

# Sense and sensitivity of double beta decay experiments

J.J. Gómez-Cadenas,<sup>a</sup> J. Martín-Albo,<sup>a</sup> M. Sorel,<sup>a</sup> P. Ferrario,<sup>a</sup>  
F. Monrabal,<sup>a</sup> J. Muñoz,<sup>a</sup> P. Novella,<sup>b</sup> A. Poves<sup>c</sup>

<sup>a</sup>Instituto de Física Corpuscular (IFIC), CSIC & Universidad de Valencia  
46980 Valencia, Spain

<sup>b</sup>Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT)  
28040 Madrid, Spain

<sup>c</sup>Departamento de Física Teórica and IFT, Universidad Autónoma de Madrid  
28049 Madrid, Spain

E-mail: [gomez@mail.cern.ch](mailto:gomez@mail.cern.ch), [justo.martin-albo@ific.uv.es](mailto:justo.martin-albo@ific.uv.es), [sorel@ific.uv.es](mailto:sorel@ific.uv.es),  
[paola.ferrario@ific.uv.es](mailto:paola.ferrario@ific.uv.es), [francesc.monrabal@ific.uv.es](mailto:francesc.monrabal@ific.uv.es), [jmunoz@ific.uv.es](mailto:jmunoz@ific.uv.es),  
[pau.novella@ciemat.es](mailto:pau.novella@ciemat.es), [alfredo.poves@uam.es](mailto:alfredo.poves@uam.es)

**Abstract.** The search for neutrinoless double beta decay is a very active field in which the number of proposals for next-generation experiments has proliferated. In this paper we attempt to address both the sense and the sensitivity of such proposals. Sensitivity comes first, by means of proposing a simple and unambiguous statistical recipe to derive the sensitivity to a putative Majorana neutrino mass,  $m_{\beta\beta}$ . In order to make sense of how the different experimental approaches compare, we apply this recipe to a selection of proposals, comparing the resulting sensitivities. We also propose a “physics-motivated range” (PMR) of the nuclear matrix elements as a unifying criterium between the different nuclear models. The expected performance of the proposals is parametrized in terms of only four numbers: energy resolution, background rate (per unit time, isotope mass and energy), detection efficiency, and  $\beta\beta$  isotope mass. For each proposal, both a reference and an optimistic scenario for the experimental performance are studied. In the reference scenario we find that all the proposals will be able to partially explore the degenerate spectrum, without fully covering it, although four of them (KamLAND-Zen, CUORE, NEXT and EXO) will approach the 50 meV boundary. In the optimistic scenario, we find that CUORE and the xenon-based proposals (KamLAND-Zen, EXO and NEXT) will explore a significant fraction of the inverse hierarchy, with NEXT covering it almost fully. For the long term future, we argue that  $^{136}\text{Xe}$ -based experiments may provide the best case for a 1-ton scale experiment, given the potentially very low backgrounds achievable and the expected scalability to large isotope masses.

**Keywords:** Double Beta Decay

**ArXiv ePrint:** [1010.5112](https://arxiv.org/abs/1010.5112)

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Rate of neutrinoless double beta decay</b>	<b>2</b>
<b>3</b>	<b>Experimental aspects</b>	<b>5</b>
<b>4</b>	<b>A unified treatment of sensitivity estimates of <math>\beta\beta 0\nu</math> experiments</b>	<b>6</b>
4.1	An ideal experiment	6
4.2	Experiments with background and the collapse of the classical limit	8
4.3	The unified approach and sensitivity of an experiment with background	10
<b>5</b>	<b>A selection of proposals</b>	<b>12</b>
5.1	Brief description of the proposals	13
5.2	Validity of background rate assumptions	16
<b>6</b>	<b>Sensitivity of the proposals</b>	<b>17</b>
6.1	Sensitivity as a function of the exposure	18
6.2	Sensitivity as a function of background rate	18
6.3	Scalability to large isotope masses	19
<b>7</b>	<b>Conclusions</b>	<b>21</b>
<b>A</b>	<b>Construction of unified-approach confidence intervals</b>	<b>22</b>

---

## 1 Introduction

Physicists invent, build and run experiments to search for new phenomena. In order to convince the funding agencies to support their research, they write proposals in which they estimate the *sensitivity* of their experiments. The definition of sensitivity is somewhat perverse: rather than promising a discovery, the proponents of a new experiment assume that *it will fail to find a signal*, and attempt to demonstrate that, in this worst-case scenario, they will exclude a larger portion of the landscape of physical parameters (such as a cross section or a lifetime) than other proposals. Loosely speaking, the sensitivity allows to compare the physics case of competing experimental approaches.

An area in which the proposals for new experiments have proliferated during the last decade is the search for neutrinoless double beta decay ( $\beta\beta 0\nu$ ) — see, for instance, [1, 2] and references therein. The detection of such a process would establish that neutrinos are Majorana particles [3, 4] (that is to say, truly neutral particles indistinguishable from their antiparticles), and a measurement of the decay rate would provide direct information on neutrino masses [5]. The observation of neutrino oscillations [6] — which implies that neutrinos have a non-zero mass, an essential condition for  $\beta\beta 0\nu$  to exist — and the possible evidence of a  $\beta\beta 0\nu$  signal in the Heidelberg-Moscow experiment [7] have boosted the interest in the double beta decay searches, and prompted a new generation of experiments with improved sensitivity. In spite of the formidable experimental challenge (or because of it) the field is rich in new ideas. Among the newcomers one finds many experimental approaches, as well

as several source isotopes. All of them claim to be capable of reaching a sensitivity to an *effective Majorana neutrino mass* (defined in Sec. 2) of at least 100 meV, and many advertise a second phase in which the sensitivity can be improved to some 20 meV.

However, considerable confusion can happen because the different proposals often use different recipes to compute the sensitivity. Also, similar sensitivity results hide sometimes more or less credible detector performance assumptions, concerning, for example, energy resolution or radiopurity. An additional complication arises when comparing the sensitivity of experimental approaches using different isotopes. In such a case, a large theoretical uncertainty exists in relating the decay rate sensitivity of a proposal to the effective Majorana neutrino mass sensitivity (see Sec. 2). And to further blur the picture, the scalability, cost, and schedules of the proposed experiments vary greatly.

In this paper we attempt to address both the sense and the sensitivity of the neutrinoless double beta decay experimental approaches currently proposed. In Section 2 we discuss the relationship between the rate of  $\beta\beta 0\nu$ , neutrino masses and nuclear theory, motivating our assumptions for the latter. Section 3 briefly reviews the experimental aspects of  $\beta\beta 0\nu$  searches. In Section 4 we discuss in detail our definition of sensitivity and show the problems that arise when the observed number of events is small compared to the expected background rate. Next, the Feldman-Cousins unified approach [8] is presented and is applied to a  $\beta\beta 0\nu$  experiment with background (we give more details of this procedure in Appendix A). Section 5 describes the main features of the experiments taken into consideration, defining a plausible range of detector performance indicators for each proposal. We try to make sense of these proposals in Section 6, where the corresponding  $m_{\beta\beta}$  sensitivities are compared as a function of exposure (kg · year) and for different scenarios. We briefly address also the question of scalability to large isotope masses.

## 2 Rate of neutrinoless double beta decay

Neutrinoless double beta decay is a very rare nuclear transition that occurs if neutrinos are massive Majorana particles [3, 4]. It involves the decay of a nucleus with  $Z$  protons into a nucleus with  $Z + 2$  protons and the same mass number  $A$ , accompanied by the emission of two electrons:  $(Z, A) \rightarrow (Z + 2, A) + 2 e^-$ . The sum of the kinetic energies of the two emitted electrons is always the same, and corresponds to the mass difference between mother and daughter nuclei,  $Q_{\beta\beta}$ . The decay violates lepton number conservation and is therefore forbidden in the Standard Model.

The simplest underlying mechanism of  $\beta\beta 0\nu$  is the virtual exchange of light ( $m_{\nu_i} < 10$  MeV) Majorana neutrinos, although, in general, any source of lepton number violation (LNV) can induce  $\beta\beta 0\nu$  and contribute to its amplitude. If we assume that the dominant LNV mechanism at low energies is the light-neutrino exchange, the half-life of  $\beta\beta 0\nu$  can be written as:

$$(T_{1/2}^{0\nu})^{-1} = G_{0\nu}(Q_{\beta\beta}, Z) \left| M_{0\nu} \right|^2 m_{\beta\beta}^2, \quad (2.1)$$

where  $G_{0\nu}(Q_{\beta\beta}, Z)$  is an exactly-calculable phase-space factor,  $|M_{0\nu}|$  is a nuclear matrix element, and  $m_{\beta\beta}$  is the effective Majorana mass of the electron neutrino:

$$m_{\beta\beta} = \left| \sum_i U_{ei}^2 m_i \right|, \quad (2.2)$$

where  $m_i$  are the neutrino mass eigenstates and  $U_{ei}$  are elements of the neutrino mixing matrix.

Neutrino oscillation experiments constrain how the effective Majorana mass in Eq. 2.2 changes with the absolute neutrino mass scale, defined as  $\min\{m_1, m_3\}$ . Only upper bounds on the absolute neutrino mass, of order 1 eV, currently exist. Also, current neutrino oscillation results cannot differentiate between two possible mass orderings, usually referred to as *normal* and *inverted* orderings. In the normal ordering case, the gap between the two lightest mass eigenstates corresponds to the small mass difference, measured by solar experiments. In this case, the effective Majorana mass can be as low as 2 meV [2]. If the spectrum is extremely hierarchical, and therefore the absolute neutrino mass can be neglected,  $m_{\beta\beta}$  can be as high as 5 meV in the normal ordering case. In the inverted ordering case, the gap between the two lightest states corresponds to the large mass difference, measured by atmospheric experiments. In this case,  $m_{\beta\beta}$  can be as low as about 15 meV [2]. The upper limit on  $m_{\beta\beta}$  if the neutrino mass can be neglected is approximately 50 meV in this case. Finally, in the particular case in which the neutrino mass differences are very small compared with its absolute scale, we speak of the *degenerate* spectrum. In this case, larger values for  $m_{\beta\beta}$  can be obtained, approximately above 50 meV.

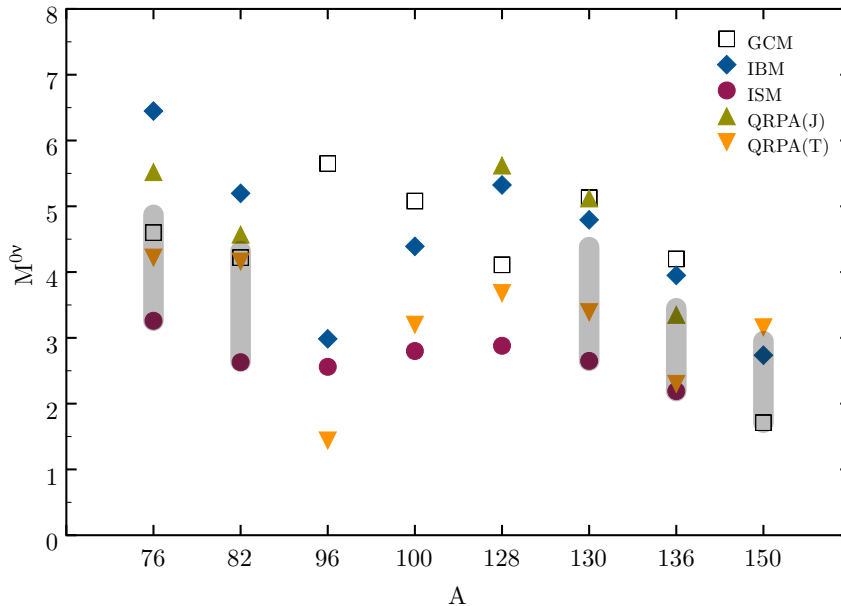
All nuclear structure effects in  $\beta\beta 0\nu$  are included in the nuclear matrix element (NME). Its knowledge is essential in order to relate the measured half-life to the neutrino masses, and therefore to compare the sensitivity and results of different experiments, as well as to predict which are the most favorable nuclides for  $\beta\beta 0\nu$  searches. Unfortunately, NMEs cannot be separately measured, and must be evaluated theoretically.

In the last few years the reliability of the calculations has greatly improved, with several techniques being used, namely: the Interacting Shell Model (ISM) [9–11]; the Quasiparticle Random Phase Approximation (QRPA) [12–14]; the Interacting Boson Model (IBM) [15]; and the Generating Coordinate Method (GCM) [16]. Figure 1 shows the most recent results of the different methods. We can see that in most cases the results of the ISM calculations are the smallest ones, while the largest ones may come from the IBM, QRPA or GCM.

Shall the differences between the different methods be treated as an uncertainty in sensitivity calculations? Should we assign an error bar to the distance between the maximum and the minimum values? This approach has, we argue, the undesirable effect of blurring the relative merits of different experimental approaches, and does not reflect the recent progress in the theoretical understanding of the treatment of nuclear matrix elements.

Each one of the major methods has some advantages and drawbacks, whose effect in the values of the NME can be sometimes explored. The clear advantage of the ISM calculations is their full treatment of the nuclear correlations, while their drawback is that they may underestimate the NMEs due to the limited number of orbits in the affordable valence spaces. It has been estimated [19] that the effect can be of the order of 25%. On the contrary, the QRPA variants, the GCM in its present form, and the IBM are bound to underestimate the multipole correlations in one or another way. As it is well established that these correlations tend to diminish the NMEs, these methods should tend to overestimate them [9, 20].

With these considerations in mind, we propose here our best estimate to NME values obtained as the central value of a physics-motivated range (PMR) of theoretical values for  $^{76}\text{Ge}$ ,  $^{82}\text{Se}$ ,  $^{130}\text{Te}$ ,  $^{136}\text{Xe}$  and  $^{150}\text{Nd}$ . In what follows we select the results of the major nuclear structure approaches which share the following common ingredients: (a) nucleon form factors of dipole shape; (b) soft short range correlations computed with the UCOM method [21]; (c) unquenched axial coupling constant  $g_A$ ; (d) higher order corrections to the nuclear current [22]; and (e) nuclear radius  $R = r_0 A^{1/3}$ , with  $r_0 = 1.2$  fm [23]. Therefore, the remaining discrepancies between the diverse approaches are solely due to the different nuclear wave



**Figure 1.** Recent NME calculations from the different techniques (GCM [16], IBM [15], ISM [10, 11], QRPA(J) [12], QRPA(T) [14, 17, 18]) with UCOM short range correlations. All the calculations use  $g_A = 1.25$ ; the IBM-2 results are multiplied by 1.18 to account for the difference between Jastrow and UCOM, and the RQRPA are multiplied by 1.1/1.2 so as to line them up with the others in their choice of  $r_0=1.2$  fm. The shaded intervals correspond to the proposed physics-motivated ranges (see text for discussion).

functions that they employ.

Let's start with the  $^{150}\text{Nd}$  case, for which no ISM value is available. The GCM calculation [16] is clearly the most sophisticated in the market from the point of view of the nuclear structure, and gives the smaller NME. The two other approaches, QRPA [18] and IBM [15], give larger and similar results; therefore we average them both to propose a NME range  $[1.71 - 2.95]$ , and therefore a NME best estimate of 2.33.

For  $^{136}\text{Xe}$  we have an ISM value which defines the lower end of the range <sup>1</sup>. For the upper one, we average the NMEs from the RQRPA calculation of the Tübingen group [14], the GCM [16], the IBM and the more recent pnQRPA result from the Jyväskylä La Plata collaboration [12]. The resulting interval is  $[2.19 - 3.45]$ , and our most probable value is 2.82.

With the same ingredients we obtain a range  $[2.65 - 4.61]$  for  $^{130}\text{Te}$ , corresponding to a best estimate of 3.63.

For  $^{82}\text{Se}$  the interval is  $[2.64 - 4.32]$  using the latest SRQRPA results [17], and the NME most probable value is 3.48.

Finally we come to  $^{76}\text{Ge}$  where we can use an extra filter, namely, to demand that the calculations must be consistent with the occupation numbers measured by J. Schiffer and collaborators [24, 25]. This leaves us with the ISM [11], the SRQRPA, [17] and the pnQRPA [12]. Averaging again the two QRPA values we obtain the interval  $[3.26 - 4.87]$ , and a best estimate of 4.07.

<sup>1</sup>We could have increased this value in view of the above discussion, but we shall refrain for doing so and use just published values.

Isotope	$W$ (g/mol)	$Q_{\beta\beta}$ (keV)	$ M_{0\nu} $	$ G_{0\nu} ^{-1}$ ( $10^{25}$ y eV <sup>2</sup> )	$T_{1/2}^{0\nu}(m_{\beta\beta} = 50 \text{ meV})$ ( $10^{27}$ y)	$N_{0\nu}/N_{0\nu}(\text{Ge})$
<sup>76</sup> Ge	75.9	2039	4.07	4.09	0.95	1.0
<sup>82</sup> Se	81.9	2996	3.48	0.93	0.26	3.3
<sup>130</sup> Te	129.9	2528	3.63	0.59	0.18	3.1
<sup>136</sup> Xe	135.9	2458	2.82	0.55	0.25	2.1
<sup>150</sup> Nd	149.9	3368	2.33	0.13	0.15	3.3

**Table 1.** Physical properties of different isotopes considered in this paper: atomic weight,  $W$  [26];  $\beta\beta$  decay  $Q$ -value [27–31]; the PMR for the nuclear matrix element ( $|M_{0\nu}|$ , see text for discussion); inverse of  $\beta\beta$  decay phase-space factor ( $|G_{0\nu}|^{-1}$ , from [32]). For illustrative purposes, we also give the  $\beta\beta 0\nu$  half-life for a fixed  $m_{\beta\beta}$  value ( $T_{1/2}^{0\nu}(m_{\beta\beta} = 50 \text{ meV})$ ), and the expected number of isotope  $\beta\beta 0\nu$  decays relative to <sup>76</sup>Ge  $\beta\beta 0\nu$  decays for a fixed isotope mass ( $N_{0\nu}/N_{0\nu}(\text{Ge})$ ).

The properties of the five isotopes relevant to our work, including our proposed values for the NMEs, are shown in Table 1. For the sake of simplicity and for the above-mentioned intended focus on directly comparing different experimental approaches, the majority of the results in this work rely only on our best estimate for NME values. On the other hand, we acknowledge that some uncertainty in NME calculations still exist, and that our recipe for estimating NME most probable values is by no means unique. For this reason, we also tabulate in Sec. 6.3 our final  $m_{\beta\beta}$  results separately for all NME calculations considered above.

### 3 Experimental aspects

The choice among  $\beta\beta$  isotopes discussed in Sec. 2 is not the only important factor to be considered when optimizing future  $\beta\beta 0\nu$  proposals. According to Table 1, one should prefer, everything else being the same, experiments based on <sup>150</sup>Nd, <sup>82</sup>Se, <sup>130</sup>Te, or even <sup>136</sup>Xe, rather than experiments based on <sup>76</sup>Ge. However, Ge-based experiments have dominated the field so far.

The reason is that  $\beta\beta 0\nu$  experiments must be designed to measure the kinetic energy of the electrons emitted in the decay. Due to the finite energy resolution of any detector,  $\beta\beta 0\nu$  events are reconstructed within a non-zero energy range centered around  $Q_{\beta\beta}$ , typically following a gaussian distribution. As will be demonstrated in Sec. 6, any background event falling in this energy range limits dramatically the sensitivity of the experiment. Good energy resolution is therefore essential. Germanium semiconductor detectors provide the best energy resolution achieved to date: in a <sup>76</sup>Ge experiment, a region able to contain most of the signal — called the *region of interest* (ROI), and often taken as 1 FWHM around  $Q_{\beta\beta}$  — would be only a few keV wide.

Unfortunately, energy resolution is not enough by itself: a continuous spectrum arising from natural decay chains can easily overwhelm the signal peak, given the enormously long decay times explored. Consequently, additional signatures to discriminate between signal and backgrounds, such as event topology or daughter ion tagging, are desirable. Also, the experiments require underground operation and a shielding to reduce external background due to cosmic rays and surrounding radioactivity, and the use of very radiopure materials. In addition, large detector masses, high  $\beta\beta$  isotope enrichment, and high  $\beta\beta$  detection efficiency are clearly desirable, given the rare nature of the process searched for. No experimental

technique scores the highest mark in all of the above, and thus different approaches are possible.

The Heidelberg-Moscow (HM) experiment [33], using high-purity germanium diodes enriched to 86% in the isotope  $^{76}\text{Ge}$ , set the most sensitive limit to date:  $T_{1/2}^{0\nu}(^{76}\text{Ge}) \geq 1.9 \times 10^{25}$  years (90% CL). The experiment accumulated a total exposure of  $71.7 \text{ kg} \cdot \text{y}$ , and achieved a background rate in the ROI of  $0.06 \text{ counts}/(\text{keV} \cdot \text{kg} \cdot \text{y})$  after pulse shape identification. The energy resolution (FWHM) at  $Q_{\beta\beta}$  was  $4.23 \pm 0.14 \text{ keV}$ . A subset of the Collaboration observed evidence for a  $\beta\beta 0\nu$  signal [7]. The claim has been severely questioned [34], but no one has been able to prove it wrong. According to it, the isotope  $^{76}\text{Ge}$  would experiment  $\beta\beta 0\nu$  decay with a lifetime of about  $1.5 \times 10^{25}$  years. Using the PMR nuclear matrix element, this corresponds to a neutrino mass of about  $0.4 \text{ eV}$ .

In this paper, we compare the  $m_{\beta\beta}$  sensitivity of different experimental approaches by parametrizing in a simple and intuitive way their detector performance, as will be discussed in Sec. 5. Before that, in Sec. 4, we discuss the statistical recipe that we will adopt to derive the experimental sensitivities.

## 4 A unified treatment of sensitivity estimates of $\beta\beta 0\nu$ experiments

### 4.1 An ideal experiment

All  $\beta\beta 0\nu$  experiments have to deal with non-negligible backgrounds, an only partially efficient  $\beta\beta 0\nu$  event selection, and more or less difficulties to extrapolate their detection technique to large masses. It is instructive, however, to imagine an ideal experiment defined by the following parameters: (a) the detector mass is 100% made of a  $\beta\beta$  source; (b) perfect energy resolution, and/or infinite radiopurity, resulting in a null background rate; (c) perfect detection efficiency; and (d) scalable to large masses at will, that is, exposure as large as desired.

Suppose that such an experiment runs for a total exposure  $Mt$ . The expected number of  $\beta\beta 0\nu$  events is given by

$$N = \log 2 \cdot \frac{N_A}{W} \cdot \varepsilon \cdot \frac{M \cdot t}{T_{1/2}^{0\nu}}, \quad (4.1)$$

where  $N_A$  is the Avogadro constant,  $W$  is the atomic weight of the  $\beta\beta$ -decaying isotope,  $\varepsilon$  is the signal detection efficiency,  $M$  is the  $\beta\beta$  isotope mass and  $t$  the data-taking time.

Different identical experiments running for the same total exposure would observe a different number of events,  $n$ , with expectation value  $\mu = N$ . The probability distribution function (pdf) of  $n$  is a Poisson distribution:

$$\text{Po}(n; \mu) = \frac{\mu^n}{n!} e^{-\mu}. \quad (4.2)$$

For example, Figure 2 (left) shows a Poisson distribution with  $\mu = 5$ . If we run a large number of identical experiments, only 17.5% will observe 5 events. In fact, the same percentage will observe 4 events, and a small but non-null number of experiments (7 in a thousand) will observe zero events. Consider now an experiment whose outcome is  $n_{\text{obs}} = 4$ . What can be inferred about the true value  $\mu$  from this measurement?

A common way to report a result on  $\mu$ , proposed by J. Neyman in 1937 [35], is to give a *confidence interval* (CI) where  $\mu$  is likely to be included. How likely an interval is to contain the true value is determined by the *confidence level* (CL), usually expressed as a

percentage. The end points of the CI, called the *confidence limits*, are defined in terms of a probability: we define the lower (upper) limit  $\mu_{\text{lo}}$  ( $\mu_{\text{up}}$ ) as the value of the parameter  $\mu$  such that, if we carry a large number of experiments following the Poisson distribution  $\text{Po}(n; \mu_{\text{lo}})$  ( $\text{Po}(n; \mu_{\text{up}})$ ), then a fraction  $\alpha$  ( $\beta$ ) of them at most will yield a number larger (smaller) than or equal to  $n_{\text{obs}}$ . Therefore, to compute the limits the following equations can be solved numerically:

$$\alpha = \sum_{n_{\text{obs}}}^{\infty} \text{Po}(n; \mu_{\text{lo}}) \quad (4.3)$$

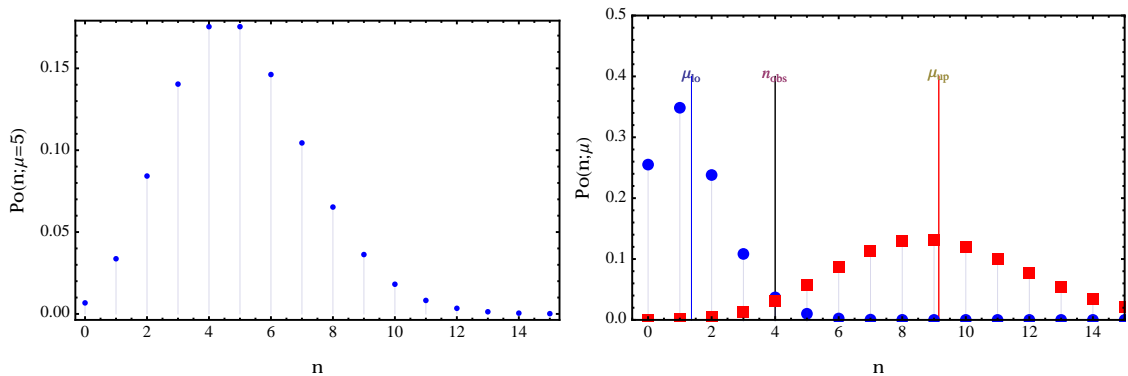
$$\beta = \sum_0^{n_{\text{obs}}} \text{Po}(n; \mu_{\text{up}}) \quad (4.4)$$

Figure 2 (right) illustrates the above definition. The lower and upper limits are computed for  $n_{\text{obs}} = 4$  and  $\alpha = \beta = 0.05$ . The filled (blue) circles represent  $\text{Po}(n; \mu_{\text{lo}} = 1.37)$ , while the filled (red) squares represent  $\text{Po}(n; \mu_{\text{up}} = 9.15)$ . Adding the filled circles corresponding to  $n < 4$  yields a probability of 0.95 (therefore, the integrated probability to find  $n \geq n_{\text{obs}} = 0.05$ ). Adding the filled squares corresponding to  $n > 4$  yields a probability of 0.95 (therefore, the integrated probability to find  $n \leq n_{\text{obs}} = 0.05$ ). Notice that our choice of a symmetric interval to *cover* the true value  $\mu = 5$  is not especially well motivated. We could have decided to set asymmetric intervals, for example,  $\alpha = 0.07$  and  $\beta = 0.03$  and would still have a 90% CL. Nevertheless, this is the most common choice for two-sided confidence intervals.

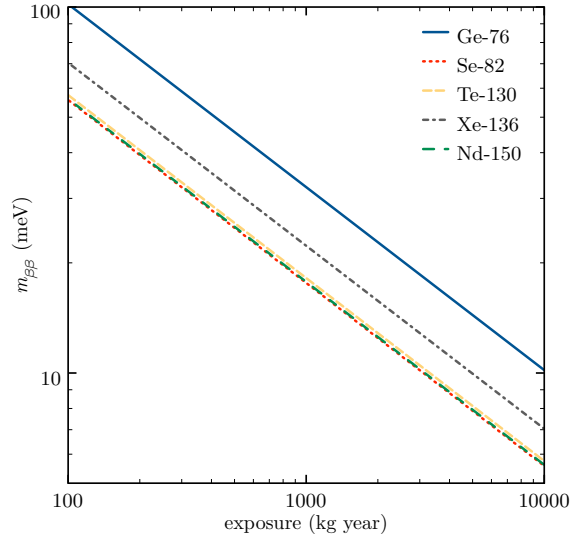
Consider now another experiment: an unlucky team runs under identical conditions than those of the previous experiment, but finds  $n_{\text{obs}} = 0$ . This is an unlikely but by no means impossible outcome, with a probability of 0.7%. What should the team report? Clearly a lower limit cannot be found, but an upper limit can still be determined by setting  $n_{\text{obs}} = 0$  in Eq. (4.4):

$$\mu_{\text{up}} = -\log(\beta) \quad (4.5)$$

For  $\beta = 0.1$  (that is, for 90% CL), Equation (4.5) yields  $\mu_{\text{up}} \simeq 2.3$ , and for  $\beta = 0.05$  (95% CL), we find  $\mu_{\text{up}} \simeq 3$ . The meaning of the upper limit is very clear. In a Poisson distribution with  $\mu = 2.3$ , one observes 10% of the times  $n = 0$ , while for  $\mu = 3.0$ , one observes 5% of the times  $n = 0$ .



**Figure 2.** Left: a Poisson distribution with  $\mu = 5$ . The probability of observing 4 events is 17.5%, which means that only that fraction of experiments will yield a number of events equal to the expectation value  $\mu$ . Right: lower ( $\mu_{\text{lo}}$ ) and upper ( $\mu_{\text{up}}$ ) limits computed for an observed number of events  $n_{\text{obs}}$  equal to 4, and  $\alpha = \beta = 0.05$ . Two Poisson distributions are shown:  $\text{Po}(n; \mu = \mu_{\text{lo}})$  and  $\text{Po}(n; \mu = \mu_{\text{up}})$ .



**Figure 3.** Sensitivity of ideal experiments at 90% CL for different  $\beta\beta$  isotopes. Since the yields are very similar, the sensitivities of  $^{82}\text{Se}$ ,  $^{130}\text{Te}$  and  $^{150}\text{Nd}$  overlap.

Therefore, an ideal experiment that observes no events after a given exposure would report a *negative* result, that is, an upper limit in the  $\beta\beta 0\nu$  decay rate  $(T_{1/2}^{0\nu})^{-1}$ , or possibly in the more relevant physical parameter  $m_{\beta\beta}$ . From Equations (2.1) and (4.1), the latter upper limit can be written as:

$$m_{\beta\beta} = K_1 \sqrt{\frac{N}{\varepsilon M t}} \quad (4.6)$$

where  $K_1$  is a constant that depends only on the isotope type. Substituting  $N$  by 2.3 (3), one gets the upper limit at 90% (95%) CL, which improves with  $\sqrt{1/Mt}$ .

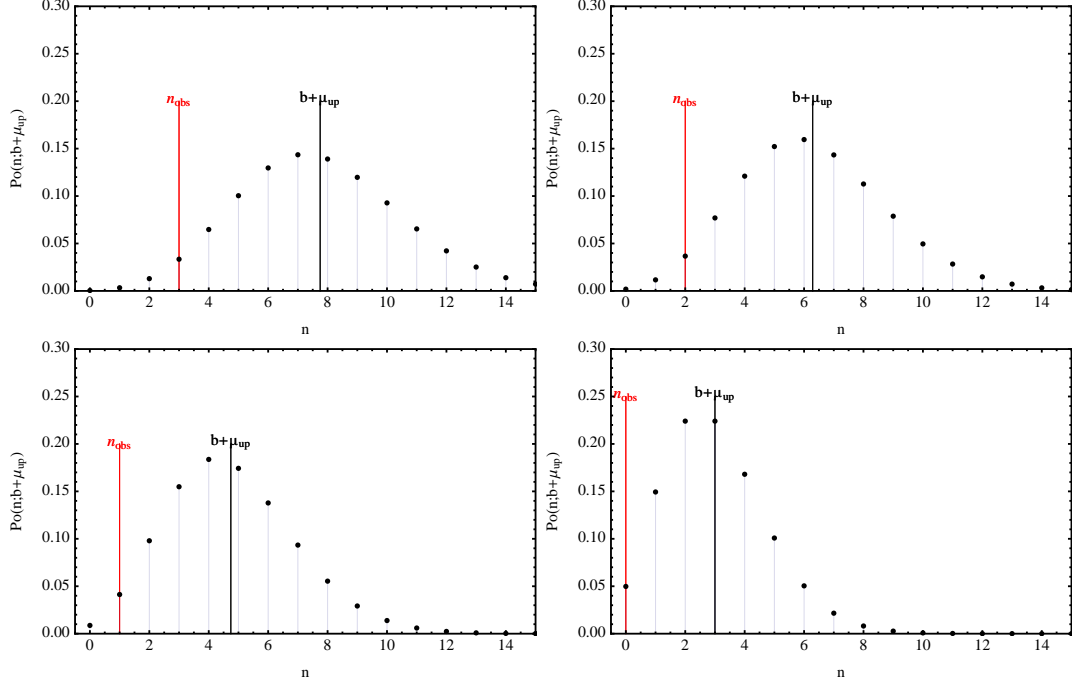
In the following, we define the sensitivity of our experiments as *the average upper limit one would get from an ensemble of experiments with the expected background and no true signal*.

For an ideal experiment, the expected background is exactly  $b = 0$ , resulting in no Poisson fluctuations in the observed number of signal plus background events for the experiments within the ensemble, which has to be always equal to zero in this case. Therefore the sensitivity of the experiment is in this case simply given by the upper limit reported by the unlucky experiment. Figure 3 shows the sensitivity of the five isotopes considered. Even ideal experiments need a large exposure ( $>10^3 \text{ kg} \cdot \text{y}$ ) to fully explore the inverse hierarchy. To start exploring the normal hierarchy, one would need of the order of  $10^4 \text{ kg} \cdot \text{y}$ , that is, perfect detectors of 1 ton running for 10 years: a truly formidable experimental challenge.

#### 4.2 Experiments with background and the collapse of the classical limit

Consider now the more realistic case of an experiment with background. Call  $b$  the expected value of the background, and assume (unrealistically) that it is known with no uncertainty. The relevant pdf will then be:

$$\text{Po}(n; \mu + b) = \frac{(\mu + b)^n}{n!} e^{-(\mu+b)}, \quad (4.7)$$



**Figure 4.** The collapse of the classical method: as the number of observed events becomes smaller, the Poisson distribution has to shift left to guarantee that the cumulative percentage to the left of  $n_{obs}$  gets to 5%. In the extreme case of  $n_{obs} = 0$ , the expectation value that yields 5% of the events in  $n = 0$ , is  $\mu + b = 3$ . Since  $b = 5$ , this forces  $\mu_{up} = -2$ , an absurd result.

where  $\mu$  is the (unknown) mean signal expectation and  $b$  is the known mean background expectation. The Poisson variable  $n$  is such that  $n = n_s + n_b$ , where the signal and background Poisson variables  $n_s$  and  $n_b$  have mean expectation values  $E[n_s] = \mu$ , and  $E[n_b] = b$ , respectively.

A priori, it appears as if we could treat the problem in exactly the same way as for the case without background. A CI can be constructed using the following equations:

$$\alpha = \sum_{n_{obs}}^{\infty} \frac{(\mu_{lo} + b)^n}{n!} e^{-(\mu_{lo} + b)} \quad (4.8)$$

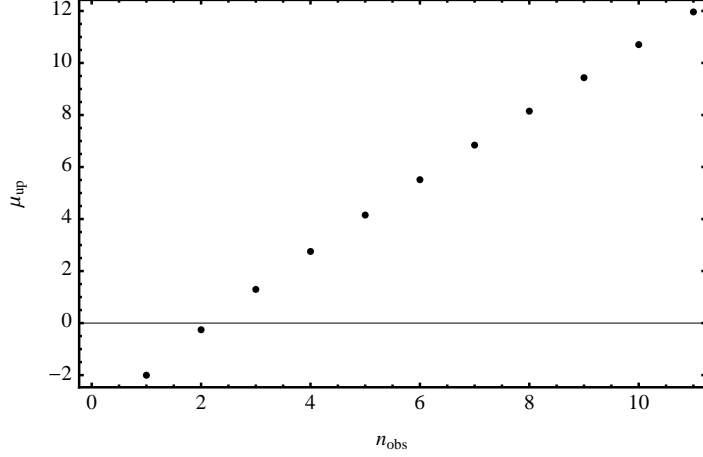
$$\beta = \sum_0^{n_{obs}} \frac{(\mu_{up} + b)^n}{n!} e^{-(\mu_{up} + b)} \quad (4.9)$$

Let us work out an explicit example. Suppose that we run an experiment in which the predicted background is  $b = 5$ , and we observe  $n_{obs} = 5$  events. Then, Equation (4.9) becomes:

$$\beta = \sum_{n=0}^5 \frac{(\mu_{up} + 5)^n}{n!} e^{-(\mu_{up} + 5)}. \quad (4.10)$$

Solving numerically this equation with  $\beta = 0.05$  (95% CL) yields an upper limit of  $\mu_{up} = 5.51$ . The lower limit is also readily computed, obtaining  $\mu_{lo} = -3.03$ . Since  $\mu$  is physically bounded to non-negative values, we set  $\mu_{lo} = 0$ .

We now repeat the procedure for  $n_{obs} = 4, 3, 2, 1$  and 0, while keeping  $b = 5$ . In this case, we obtain  $\mu_{up} = 4.15, 2.75, 1.29, -0.25$  and  $-2$ , respectively. The reason for the decrease in  $\mu_{up}$  as  $n_{obs}$  decreases can be understood by observing Fig. 4. As the number of observed events becomes smaller, the Poisson distribution has to shift left to guarantee that



**Figure 5.** Upper limit  $\mu_{\text{up}}$  as a function of the number of observed events  $n_{\text{obs}}$ , estimated solving Eq. (4.9) for a background prediction  $b = 5$ . The limit does not make sense when the observed number of events is small compared with the expected background ( $n_{\text{obs}} \ll b$ ).

the cumulative percentage to the left of  $n_{\text{obs}}$  gets to 5%. The situation is summarized in Fig. 5: due to the discreteness of the Poisson distribution, the “classical upper limit” method fails to provide meaningful results (i.e., it yields negative upper limits) when the observed number of events is small compared with the expected number of background events.

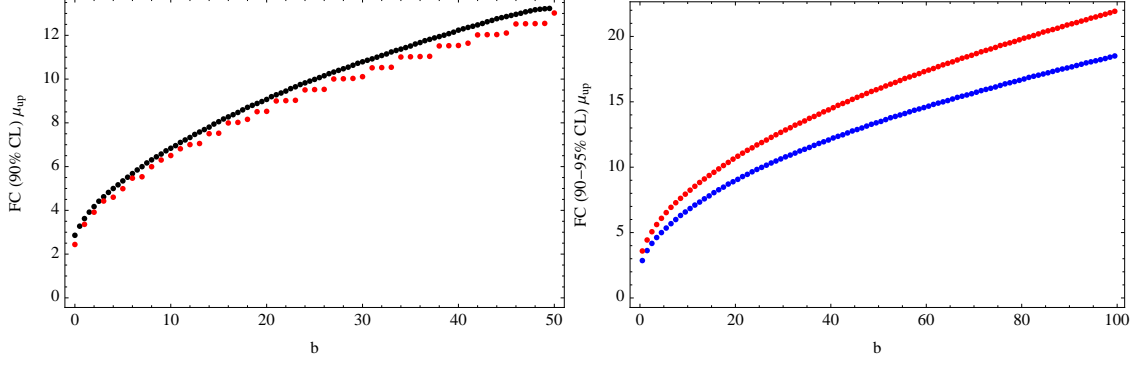
### 4.3 The unified approach and sensitivity of an experiment with background

The collapse of the classical limit was solved in 1997 by G. Feldman and R. Cousins. In their best-seller paper, “Unified approach to the classical statistical analysis of small signals” [8], they introduce a method for the construction of confidence intervals using an ordering principle based on likelihood ratios. The procedure avoids the collapse described in Sec. 4.2, unifying the treatment of upper confidence limits for null results and two-sided confidence limits for non-null results. Since its publication, the unified approach method has become the “de facto” standard frequentist approach to compute confidence intervals. We give some details of the method in the Appendix A.

We revisit now the concept of sensitivity in the case of an experiment with background. Recall that by sensitivity we mean the average upper limit that would be obtained by an ensemble of identical replicas of such an experiment, each one with the same mean expected background and no true signal. Let us now translate the above definition into a mathematical formula. We define  $\mathcal{U}(n|b)$  as the function yielding the (unified approach) upper limit (at the desired CL) for a given observation  $n$  and a mean predicted background level  $b$ . Values for  $\mathcal{U}(n|b)$  are reported in tabular form in [8] for several CL values. Given that the variable  $n$  follows a Poisson pdf,  $\text{Po}(n|b) \equiv \text{Po}(n; b)$ , then, according to our definition, the sensitivity  $\mathcal{S}(b)$  is given by:

$$\mathcal{S}(b) \equiv \text{E}[\mathcal{U}(n|b)] = \sum_{n=0}^{\infty} \text{Po}(n|b) \mathcal{U}(n|b). \quad (4.11)$$

This equation reads as follows: the sensitivity  $\mathcal{S}(b)$  of an experiment expecting  $b$  events of background is obtained by averaging the upper limits obtained using the unified approach ( $\mathcal{U}(n|b)$ ) with the likelihood of the individual observations ( $\text{Po}(n|b)$ ).



**Figure 6.** Left: 90% CL sensitivity curve as a function of mean background prediction  $b$  as obtained according to Eq. (4.11) (black dots) or by simply assuming  $n = b$  (red dots). Right: sensitivity curve as a function of  $b$  for 90% and 95% CL.

Let us work out an example. Assume that an experimental team proposes a 100 kg germanium detector to run underground for 10 years. The team claims that the expected background rate in the experiment is  $c = 0.001$  counts/(keV · kg · y). Their energy ROI has a width of 5 keV. The predicted background is then  $b = 0.001 \times 100 \times 5 \times 10 = 5$  events. If we set  $\mu = 0$  (no signal) and  $b = 5$  in Eq. (4.7), we obtain the same distribution as for the case ( $\mu = 5$ ,  $b = 0$ ) described in Sec. 4.1. If we throw many experiments following  $\text{Po}(n; \mu = 5)$ , the fraction of the times that we will obtain  $n_{\text{obs}} = 0, 1, 2, 3 \dots$ , is therefore described by the pdf shown in the top panel of Fig. 2. Consider for instance the case  $n = 5$ . The upper limit at 90% CL following the prescription of the unified approach is  $\mathcal{U}(5|5) = 4.99$  (see [8]); we also know that  $\text{Po}(5|5) = 0.175$  (see Fig. 2). Analogously, for  $n = 4$ ,  $\mathcal{U}(4|5) = 3.66$  and  $\text{Po}(4|5) = 0.175$ ; for  $n = 3$ ,  $\mathcal{U}(3|5) = 2.78$  and  $\text{Po}(3|5) = 0.14$ ; and so forth. The sensitivity of the experiment, computed as the average of all the experiments, would then be:

$$\begin{aligned} \mathcal{S}(5) &= \sum_{n=0}^{\infty} \text{Po}(n|5) \mathcal{U}(n|5) \\ &= 4.99 \times 0.175 + 3.66 \times 0.175 + \dots = 5.17 \end{aligned}$$

This value,  $\mathcal{S}(5) = 5.17$  (90% CL) is different than the one sometimes used as upper limit,  $\mathcal{U}(5|5) = 4.99$ . As illustrated in the left panel of Fig. 6, the difference between the  $\mathcal{S}(b)$  and  $\mathcal{U}(b|b)$  sensitivity estimates is small but significant for all  $b$ . The latter approach is not strictly correct, introduces fluctuations, and underestimates the sensitivity limit for all  $b$ . The right panel in Fig. 6 compares the 90% with the 95% CL sensitivity curves obtained with our sensitivity procedure.

We should note that, in the large background approximation, the sensitivity curve as a function of  $b$  follows the expected classical limit:

$$\mu_{\text{up}} \equiv \mathcal{S}(b) \simeq \alpha \cdot \sqrt{b}, \text{ for large } b, \quad (4.12)$$

where  $\alpha = 1.64$  (1.96) at 90% (95%) CL.

If we substitute  $N$  in Eq. (4.6) by the upper limit obtained using the unified approach, then, in the limit of large background:

$$m_{\beta\beta} = K_2 \sqrt{\frac{b^{1/2}}{\varepsilon M t}} \quad (4.13)$$

where  $K_2$  is a constant depending on the isotope. If the background  $b$  is proportional to the exposure  $Mt$  and to the width  $\Delta E$  of the ROI:

$$b = c \cdot Mt \cdot \Delta E \quad (4.14)$$

with the background rate  $c$ , in counts/(keV · kg · y), constant across the energy ROI, then:

$$m_{\beta\beta} = K_2 \sqrt{1/\varepsilon} \left( \frac{c \Delta E}{Mt} \right)^{1/4} \quad (4.15)$$

That is, in the limit of large background, the sensitivity to  $m_{\beta\beta}$  improves very slowly with exposure, as  $(Mt)^{-\frac{1}{4}}$ .

In this paper, we compare the sensitivity of several experiments using the unified approach to compute upper limits after Equation (4.11). As it was the case for our choice of the NME, this is a conservative but robust option. It is conservative because we do not use all the information available to compute upper limits (such as the potentially different energy distributions for the signal and the backgrounds in the ROI). It is robust, however, since the unified approach ensures coverage, unlike, for example, the profile likelihood method in the low background limit. Furthermore, it does not require the precise knowledge of the pdfs of signal and backgrounds in the ROI for the different experiments. While other methods can be used and many have been proposed in the literature (see, for example, [36]), we believe that the use of the unified approach allows a simple (and therefore easily reproducible) calculation of sensitivities. As discussed in Sec. 3 and motivated here in the large exposure limit (but valid for all exposures), proponents of an experiment simply need to provide three parameters describing the detector performance and the isotope to be used, in order to allow the derivation of the  $m_{\beta\beta}$  sensitivity for a given exposure. The three parameters are: (a) the FWHM energy resolution  $\Delta E$ , defining the energy ROI (of, say, 1 FWHM around  $Q_{\beta\beta}$ ) for the  $\beta\beta 0\nu$  search; (b) the background rate  $c$  per unit of  $\beta\beta$  isotope mass, energy and time; (c) the  $\beta\beta$  detection efficiency  $\varepsilon$ . With this information, the sensitivity can be computed using Equation (4.11).

## 5 A selection of proposals

Many double beta decay experiments have been proposed in the last decade. A recent review [2] quotes 14 projects in different stages of development and using a variety of detection techniques and isotopes. In this paper we compare the relative merits of 7 of these proposals, chosen as representatives of four different experimental approaches. For each of the 7 experiments considered, we report in Table 2 the  $\beta\beta$  isotope that will (likely) be used, together with the three parameters discussed in Sec. 4.3 that describe the detector performance. It is possible to find all these figures in the literature, although the level of detail in the discussion and the uncertainties associated with them vary greatly from case to case.

Two scenarios have been considered regarding the background rates expected in the experiments: the *reference* (R) scenario, which implies in most cases an improvement of at least one order of magnitude over the state of the art represented by Heidelberg-Moscow, and the *optimistic* (O) scenario, which projects a further improvement with respect to the R scenario. In order to evaluate the scalability of the proposals to large masses (see Sec. 6.3), we give also a reference and an optimistic assumption for the  $\beta\beta$  isotope mass.

In the following, Sec. 5.1, a brief description of the proposals is given. The validity of the background rate assumptions, in light of the results achieved so far, is discussed in Sec. 5.2.

Experiment	Isotope	Resolution (% at $Q_{\beta\beta}$ )	Efficiency		Background rate ( $10^{-3}$ cts/(keV kg y))	Mass (kg)
CUORE	$^{130}\text{Te}$	0.18	0.80	R	40	200
				O	1	400
EXO	$^{136}\text{Xe}$	3.3	0.70	R	1	160
				O	0.5	1000
GERDA	$^{76}\text{Ge}$	0.16	0.80	R	10	15
				O	1	35
KamLAND-Zen	$^{136}\text{Xe}$	9.5	0.80	R	0.5	360
				O	0.1	1000
NEXT	$^{136}\text{Xe}$	0.7	0.30	R	0.2	90
				O	0.06	1000
SNO+	$^{150}\text{Nd}$	6.5	0.50	R	10	50
				O	1	500
SuperNEMO	$^{82}\text{Se}$	4.0	0.30	R	0.4	7
				O	0.06	100

**Table 2.** Proposals considered in the  $m_{\beta\beta}$  sensitivity comparison. For each proposal, the isotope that will (likely) be used, together with estimates for detector performance parameters — FWHM energy resolution, detection efficiency and background rate per unit of energy, time and  $\beta\beta$  isotope mass — are given. The efficiencies indicated here do not include the efficiency loss due to the 1 FWHM energy cut, common to all proposals. The background estimates and the  $\beta\beta$  source mass include both a reference (R) and an optimistic (O) scenario.

### 5.1 Brief description of the proposals

The first experimental approach we consider is the **high-resolution calorimetry**. These are detectors characterized by excellent energy resolution. They have also the advantages of simplicity and compactness. The classic germanium diodes and the bolometers fall in this category.

Two experiments, **GERDA** [37–40] and **MAJORANA** [41–43], will search for  $\beta\beta 0\nu$  in  $^{76}\text{Ge}$  using arrays of high-purity germanium detectors. This is a well-established technique that offers outstanding energy resolution (better than 0.2% FWHM at the  $Q$ -value) and high efficiency ( $\sim 0.80$ ). In its first phase GERDA expects a background rate of the order of  $10^{-2}$  counts/(keV · kg · y). The Collaboration aims to improve this by an additional order of magnitude,  $10^{-3}$  counts/(keV · kg · y) in the ROI, during its second phase. For its part, MAJORANA anticipates a background rate of  $10^{-3}$  counts/(ROI·kg·y), that is, about 4 times less than the second phase of GERDA. For the sake of concreteness, we choose GERDA to represent the current germanium-based proposals, assigning its first phase —  $10^{-2}$  counts/(keV · kg · y) and 15 kg of isotope mass — to the R scenario and its second phase —  $10^{-3}$  counts/(keV · kg · y) and 35 kg — to the O scenario<sup>2</sup>.

<sup>2</sup>The GERDA and MAJORANA Collaborations plan to merge in the future for a  $^{76}\text{Ge}$  experiment aiming for a ton-scale isotope mass and for a  $10^{-4}$  counts/(keV · kg · y) background rate. Such third phase for future  $^{76}\text{Ge}$  detectors is considered beyond the scope of this work, focusing on experimental proposals for the coming decade.

**CUORE** [44–46] is an array of  $\text{TeO}_2$  bolometers. Because  $^{130}\text{Te}$  has a large natural abundance ( $\sim 34\%$ ), the need for enrichment is less important than in other isotopes. CUORE can collect a large mass of isotope ( $\sim 200$  kg for a total detector mass of about 740 kg). The advantages of the technique are similar to those of germanium experiments with about the same energy resolution and efficiency for the signal. Studies have already demonstrated background rates at the  $0.2$  counts/(keV  $\cdot$  kg  $\cdot$  y) level. We assign  $4 \times 10^{-2}$  counts/(keV  $\cdot$  kg  $\cdot$  y) and  $10^{-3}$  counts/(keV  $\cdot$  kg  $\cdot$  y) to the reference and optimistic scenarios respectively. Concerning the mass, we assume the full 200 kg of CUORE in the R scenario, and a possible doubling (by enriching the target to some 70%) of the mass for the O scenario.

The second approach considered is **xenon time-projection chambers**. Xenon is a suitable detection medium, providing both scintillation and ionization signals. It has a  $\beta\beta$ -decaying isotope,  $^{136}\text{Xe}$ , with a natural abundance of about 10%. Compared to other  $\beta\beta$  sources, xenon is easy (thus relatively cheap) to enrich in the candidate isotope. It has no other long-lived radioactive isotopes, and no spallation products constitute a background. There are two possibilities for a xenon TPC: a cryogenic liquid xenon TPC (LXe), or a (high pressure) gas chamber (HPXe). We consider examples of both options here.

The Enriched Xenon Observatory (**EXO**) [47, 48] will search for  $\beta\beta 0\nu$  decay in  $^{136}\text{Xe}$  using a 200-kg liquid-xenon (enriched at 80% in  $^{136}\text{Xe}$ ) TPC during its first phase. The use of liquefied xenon results in a relatively modest energy resolution, 3.3% FWHM at  $Q_{\beta\beta}$  [49] (using the anti-correlation between ionization and scintillation). The signal detection efficiency is estimated by the Collaboration to be about 70%. Background rates of order  $10^{-3}$  counts/(keV  $\cdot$  kg  $\cdot$  y) are expected in EXO-200. The improvement with respect to the high-resolution calorimeters comes from the event topological information, that allows the rejection of superficial backgrounds and multi-hit events, so that only energetic gammas from  $^{208}\text{Tl}$  and  $^{214}\text{Bi}$  constitute a significant source of background. The optimistic scenario,  $5 \times 10^{-4}$  counts/(keV  $\cdot$  kg  $\cdot$  y), projects a further improvement in the detector radiopurity. The ultimate goal of the EXO Collaboration is to develop the so-called *barium tagging* [50]. This technique would allow the detection of the ion product of the  $^{136}\text{Xe}$  decay, and thus eliminate all backgrounds but the intrinsic  $\beta\beta 2\nu$ . For this study, however, we have not considered its benefits. Concerning the mass, we take 200 kg of Xe (160 kg of  $^{136}\text{Xe}$ ) for the reference scenario, and assume one ton of isotope for the optimistic scenario.

The **NEXT** Collaboration is building a 100 kg high-pressure gaseous xenon (enriched at 90% in  $^{136}\text{Xe}$ ) TPC [51]. The experiment aims to take advantage of both good energy resolution ( $\lesssim 1\%$  FWHM at  $Q_{\beta\beta}$ ) and the presence of a  $\beta\beta 0\nu$  topological signature for further background suppression. As a result, the background rate is expected to be one of the lowest among all the proposals considered here:  $2 \times 10^{-4}$  counts/(keV  $\cdot$  kg  $\cdot$  y) in the reference scenario, with a further improvement by a factor of 3 in the optimistic scenario. The low background rate, however, comes at the expense of a relatively inefficient signal detection:  $\sim 30\%$ . In order to reach its ambitious goals, the NEXT Collaboration plans to rely on electroluminescence to amplify the ionization signal, using two separate photo-detection schemes for an optimal measurement of both calorimetry and tracking [52]. Like in the case of EXO we assume 100 kg of xenon (90 kg of  $\beta\beta$  isotope) for the reference scenario and 1 ton for the optimistic scenario.

The third category of experiments is **large self-shielding calorimetry**. These are large detectors in which the source, dissolved in liquid scintillator, is surrounded by a large buffer volume. The advantages of the approach are the self-shielding against external back-

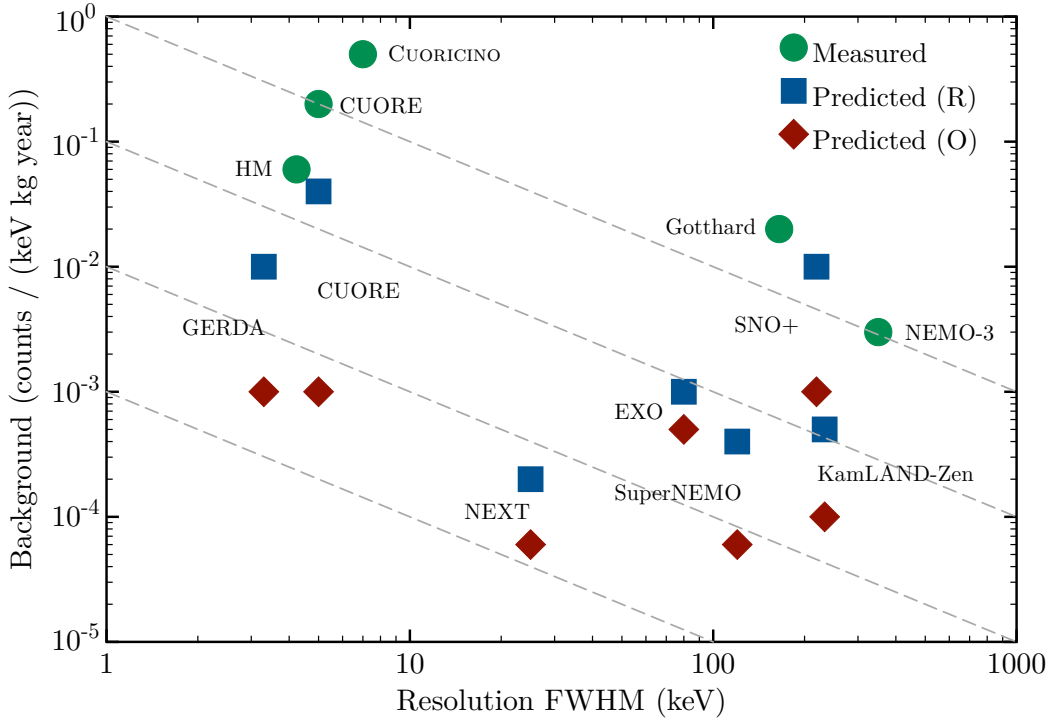
grounds and the scalability, since it appears possible to dissolve large isotope masses. The main disadvantage is poor energy resolution.

**SNO+** [53–55] proposes to fill the Sudbury Neutrino Observatory (SNO) with ultra-pure liquid scintillator. A mass of several tens of kilograms of  $\beta\beta$ -decaying material can be added to the experiment by dissolving a neodymium salt in the scintillator. The natural abundance in the  $^{150}\text{Nd}$  isotope is 5.6%. Given the liquid scintillator light yield and photocathode coverage of the experiment, a modest energy resolution performance (about 6% FWHM at  $Q_{\beta\beta}$ ) is expected. External backgrounds can be rejected with a relatively tight fiducial volume selection, cutting about  $\sim 50\%$  of the signal. A background rate similar to that of GERDA in its first phase is expected ( $\sim 10^{-2}$  counts/(keV  $\cdot$  kg  $\cdot$  y) in the R scenario, improving by a factor 10 in the O scenario. Concerning the isotope mass, SNO+ has the disadvantage of using natural neodymium, resulting in about 50 kg of isotope for the reference scenario. The optimistic scenario assumes that enriched neodymium can be use, resulting in a factor 10 increase of the mass.

The **KamLAND-Zen** [56, 57] experiment plans to dissolve 400 kg of  $^{136}\text{Xe}$  in the liquid scintillator of KamLAND in the first phase of the experiment, and up to 1 ton in a projected second phase. Xenon is relatively easy to dissolve (with a mass fraction of more than 3% being possible) and also easy to extract. The major modification to the existing KamLAND experiment is the construction of an inner, very radiopure ( $10^{-12}$  g/g of  $^{238}\text{U}$  and  $^{232}\text{Th}$ ) and very transparent balloon to hold the dissolved xenon. The balloon, 1.7 meters in radius, would be shielded from external backgrounds by a large, very radio-pure liquid scintillator volume. While the energy resolution at  $Q_{\beta\beta}$  (about 10%) is inferior to that of SNO+, the detection efficiency is much better (80%) due to its double envelope. The double envelope design is also responsible for the low expected background rate,  $2 \times 10^{-4}$  counts/(keV  $\cdot$  kg  $\cdot$  y) in the R scenario, with a further improvement by a factor of 2 in the O scenario.

The fourth and last category is that of **Tracko-Calo experiments**, where foils of  $\beta\beta$  source are surrounded by a tracking detector that provides a direct detection of the two electron tracks emitted in the decay. The quintessential example of this technique is the SuperNEMO experiment.

**SuperNEMO** [58–60] is a series of modules, each one consisting of a tracker and a calorimeter that surround a thin foil of the  $\beta\beta$  isotope. In SuperNEMO the target will likely be  $^{82}\text{Se}$ , although other isotopes such as  $^{150}\text{Nd}$  or  $^{48}\text{Ca}$  are also being considered. The mass of the target is limited to a few kg (typically 5 to 7) by the need to build it foil-like, and to minimize multiple scattering and energy loss. The tracker and calorimeter can record the trajectory of the charged particles and measure their energies independently. As shown by the successful NEMO-3 experiment [61–63], this technique, which exploits maximally the topological signature of the events, leads to excellent background rejection. Moreover, Tracko-Calo experiments allow for the determination of the  $\beta\beta 0\nu$  decay mechanism as the individual energies and trajectories of both electrons are measured [64]. In exchange, the selection efficiency is — like in the case of the NEXT experiment — relatively low (about 30%), and the resolution rather modest (4% FWHM at  $Q_{\beta\beta}$ ). This technique is very hard to extrapolate to large masses due to the size, complexity and cost of each module. For this reason, we consider that the target mass of the SuperNEMO collaboration (100 kg, or about 20 modules) will only be achieved in the optimistic scenario. For the reference scenario we consider 7 kg. This is the isotope mass for the *demonstrator* module that the Collaboration expects to commission in 2013.



**Figure 7.** Background rate in the ROI plotted versus the energy resolution (FWHM) for different past and present experiments. The dashed lines delimit bands where the experiments falling inside have background rates per unit of exposure — counts/(kg·year) — of the same order of magnitude. (Figure adapted from [65].)

## 5.2 Validity of background rate assumptions

How justified are the assumptions of the different experiments regarding their expected level of background suppression? Figure 7 offers some clues. It shows the background rate in the ROI — in counts/(keV · kg · y) — versus the energy resolution (FWHM) of several past and present experiments. The (green) circles correspond to measured data, while the (blue) squares and (red) diamonds correspond, respectively, to the R and O background assumptions of the experiments discussed in the previous section. The dashed lines delimit bands where the experiments falling inside have background rates per unit of exposure — counts/(kg·y) — of the same order of magnitude.

The best background rates obtained so far are rather modest if one considers the ambitions of the collaborations. Even in the case that we call *reference*, achieving the expected performance poses a major experimental challenge.

The GERDA experiment expects to reduce the background rate achieved in Heidelberg-Moscow — 0.06 counts/(keV · kg · y) after pulse-shape discrimination — by at least a factor of 6 (phase I), and up to a factor of 60 (phase II). This improvement comes from a better shielding and improved radiopurity, and from active background suppression techniques such as rejection of multi-site depositions [66–68].

CUORE aims at achieving a background rate of about  $4 \times 10^{-2}$  counts/(keV · kg · y). A previous effort, the CUORICINO detector [69], a tower of bolometers that accumulated an exposure of 11.83 kg · years, measured a background rate of 0.6 counts/(keV · kg · y).

More recent tests within the CUORE Collaboration have measured about 0.2 counts/(keV · kg · y) [45]. Further improvements in the background rate would require, probably, active suppression methods such as those proposed by the LUCIFER project [2]. Such techniques may permit a reduction of the background rate down to  $10^{-3}$  counts/(keV · kg · y) or better.

The xenon time-projection chambers have a precedent in the Gotthard experiment [70], a small xenon TPC ( $\sim 5$  kg) operated at 5 bar that obtained a background rate in the ROI of about 0.02 counts/(keV · kg · y) using event topological information to reject backgrounds. The NEXT experiment expects a factor of 100 improvement over Gotthard due to an upgraded detection technique: the Gotthard TPC suffered of modest energy resolution — 6.6% FWHM at the  $Q$ -value — probably due to the addition of methane to the xenon (in order to increase the drift velocity and to suppress diffusion), which quenched the scintillation and ionization signals. NEXT aims at achieving energy resolution better than 1% FWHM at  $Q_{\beta\beta}$ , enough to separate the  $\beta\beta 0\nu$  signal peak from the Compton spectrum of  $^{208}\text{Tl}$ , the main background source in Gotthard. Additionally, NEXT will measure the  $t_0$  of each event using the primary scintillation signal, a handle that Gotthard lacked. This permits the definition of a fiducial volume that eliminates any charged track emanating from detector surfaces. EXO projects a background level of  $10^{-3}$  counts/(keV · kg · y), thanks to their good self-shielding. This background rate is a factor 5 higher than that of NEXT and compatible with the availability of a topological signature in the pressurized gas, not present in the liquid phase.

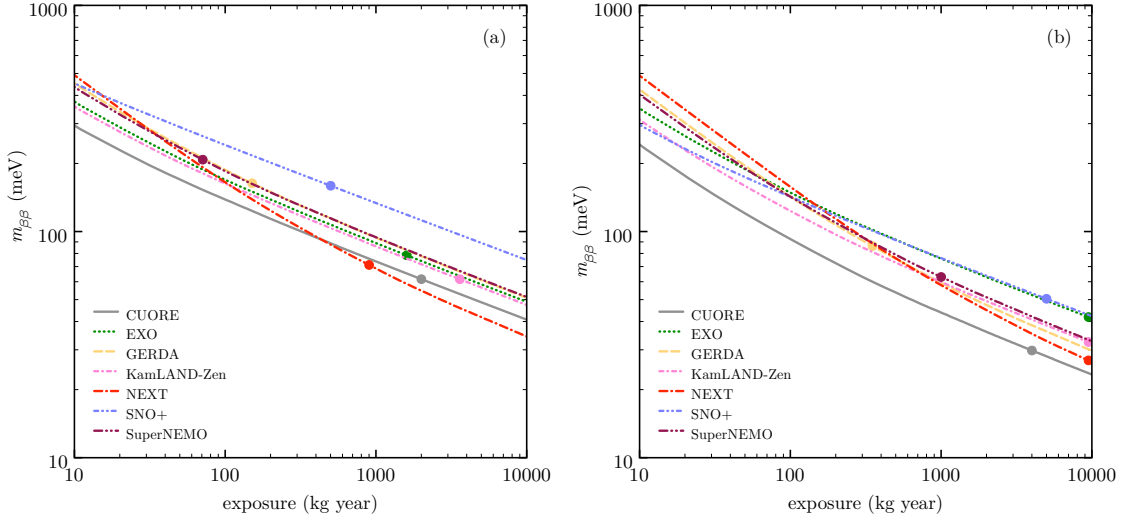
SNO+ projects a background level similar to that of EXO, also due to good self-shielding. KamLAND-Zen, however, expects a factor of 5 lower rate than both EXO and SNO+ and at the same level of NEXT and SuperNEMO, in spite of lacking a topological signature. The advantage over SNO+ is the inner balloon that allows them to dissolve the xenon in a smaller volume of liquid scintillator. This reduces the fractional background due to the scintillator itself and improves the shielding.

The NEMO-3 experiment has achieved a spectacular background rate of  $3 \times 10^{-3}$  counts/(keV · kg · y) using the topological signature of the events to discriminate signal from background. SuperNEMO foresees an improvement of up to a factor of 60 in the background rate with respect to its predecessor. This implies challenging requirements in the radiopurity of the source foil and in the level of radon inside the detector (the two main sources of background in NEMO-3), since the experimental technique is essentially the same.

## 6 Sensitivity of the proposals

We now turn to the results concerning the  $m_{\beta\beta}$  sensitivity, as defined in Sec. 4, for the various  $\beta\beta 0\nu$  proposals discussed in Sec. 5. In Sec. 6.1, we show our results for the  $m_{\beta\beta}$  sensitivity of the proposals as a function of exposure, for both our reference and optimistic scenarios. Section 6.2 illustrates how the sensitivity of the proposals is affected by changes in the background rate assumptions. We comment on the scalability to large isotope masses of the various proposals, and on their plausible  $m_{\beta\beta}$  sensitivity after a common data-taking period of 10 years, in Sec. 6.3. The efficiency of each experiment, as well as the overall efficiency for a ROI of 1 FWHM, is included in all the results.

A caveat is in order. Although we believe that our analysis is robust and simple, it also has some limitations. On the one hand, the quoted sensitivities could still improve somewhat if an energy-dependent analysis is performed. However, this requires a detailed understanding of the energy distribution of the various backgrounds in each experiment, and such information is hard to find in the existing literature. On the other hand, adding system-



**Figure 8.** The  $m_{\beta\beta}$  sensitivity (at 90% CL) as a function of exposure of the seven different  $\beta\beta 0\nu$  proposals considered. For each proposal, two detector performance scenarios are shown: (a) reference case, (b) optimistic case. For illustrative purposes, the filled circles indicate 10 years of run-time according to the reference and optimistic isotope mass assumptions in Table 2.

atic errors in both the estimation of the background and the calculation of the efficiency will partially spoil the sensitivity. This last point may severely affect the sensitivity results, but it is very difficult to implement in a general study like ours, since the published information concerning systematic errors of the different proposals is very scarce.

### 6.1 Sensitivity as a function of the exposure

Figure 8 shows the  $m_{\beta\beta}$  sensitivity (at 90% CL) as a function of exposure for the seven  $\beta\beta 0\nu$  proposals reviewed. We consider here not only the impact of isotope choice, efficiency and energy resolution, but also the effect of the expected background rate in the R and O scenarios. For the sake of simplicity, let us focus here on a fixed exposure of  $10^3 \text{ kg} \cdot \text{y}$  for all proposals.

In the reference scenario and for the above-mentioned exposure, NEXT and CUORE have the best sensitivities, reaching 69 and 74 meV at 90% CL, respectively. KamLAND-Zen, EXO, GERDA and SuperNEMO follow, with sensitivities around 90 meV. SNO+ reaches a sensitivity 134 meV. Taking into account that these numbers represent a limit of maximum exposure for most of the proposals, 1 ton-year, it follows that the “next-generation” experiments will, at best, explore the degenerate spectrum.

In the optimistic scenario, the lower background regime for all experiments allows to obtain significantly better sensitivities for the same  $10^3 \text{ kg} \cdot \text{y}$  exposure. In this case, the best sensitivity is obtained by CUORE (44 meV at 90% CL). NEXT, KamLAND-Zen, GERDA and SuperNEMO follow, with sensitivities around 60 meV. EXO and SNO+ reach sensitivities around 75 meV.

### 6.2 Sensitivity as a function of background rate

Table 3 gives the sensitivity of all the proposals (for an exposure of 1 ton-year) and for five different background rate assumptions. We can observe the following trends:

Experiment	$m_{\beta\beta}$ (meV)				
	$c = 10^{-1}$	$c = 10^{-2}$	$c = 10^{-3}$	$c = 10^{-4}$	$c = 10^{-5}$
CUORE	93	54	34	25	24
EXO	258	147	85	51	37
GERDA	155	91	58	45	42
KamLAND-Zen	314	178	101	60	39
NEXT	269	154	91	61	50
SNO+	237	134	76	45	30
SuperNEMO	354	201	115	69	48

**Table 3.** Large exposure  $m_{\beta\beta}$  sensitivity (for  $Mt = 10^3$  kg · y and at 90% CL) of the different  $\beta\beta 0\nu$  experiments in terms of the background rate (expressed in counts/(keV · kg · y)).

- In the background rate characterizing past  $\beta\beta 0\nu$  experiments,  $c \sim 10^{-1}$  counts/(keV · kg · y), CUORE would be the experiment with best sensitivity, 93 meV at 90% CL.
- With a one order of magnitude better background rate suppression,  $c = 10^{-2}$  counts/(keV · kg · y), GERDA would reach 91 meV at 90% CL and CUORE would almost fully explore the degenerate spectrum (down to 54 meV at 90% CL).
- For background rates at the  $10^{-3}$  counts/(keV · kg · y) level, CUORE would partially probe the inverted hierarchy ( $m_{\beta\beta} \sim 34$  meV), GERDA would almost fully explore the degenerate spectrum, while EXO, NEXT and SNO+, would have a sensitivity better than 100 meV.
- For even lower background rates,  $c = 10^{-4}$  counts/(keV · kg · y), NEXT, EXO, KamLAND-Zen and SuperNEMO would almost fully cover the degenerate spectrum; GERDA would probe a small part of the inverse hierarchy; and CUORE would cover it almost completely.
- In the very low background rate limit of  $c = 10^{-5}$  counts/(keV · kg · y), one can see that the improvement in sensitivity of both CUORE and GERDA is very small, indicating that these experiments have reached the “background-free” regime for the considered exposure. NEXT, EXO, KamLAND-Zen and SuperNEMO, instead, keep improving their sensitivity, and explore partially the inverse hierarchy.

### 6.3 Scalability to large isotope masses

Having considered  $m_{\beta\beta}$  sensitivity as a function of exposure in kg · y, we consider now the mass scalability of the different proposals.

Germanium experiments such as GERDA and MAJORANA are easier to scale, given their compactness. While masses of the order of 100 kg appear possible, it may be difficult, in particular for economical reasons, to go much beyond that. CUORE aims already at a large mass, and can, presumably, scale up at least by a factor two if enrichment at large scale is feasible and the backgrounds can be kept under control.

It will be very difficult to scale up SuperNEMO. One is faced here with the difficulty of replicating modules (about 20 are needed), high cost, and the need of a very large underground volume to host the modules and their shielding. The technique is clearly unsuitable for

Experiment		$m_{\beta\beta}$ (meV)					
		PMR	GCM	IBM	ISM	QRPA(J)	QRPA(T)
CUORE	R	<b>62</b>	44	47	84	44	66
	O	<b>30</b>	21	23	41	21	32
EXO	R	<b>79</b>	53	56	101	66	96
	O	<b>41</b>	28	29	53	35	51
GERDA	R	<b>163</b>	144	103	204	120	157
	O	<b>86</b>	76	54	108	64	83
KamLAND-Zen	R	<b>62</b>	41	44	79	52	75
	O	<b>32</b>	22	23	41	27	39
NEXT	R	<b>71</b>	48	51	92	60	87
	O	<b>27</b>	18	19	34	22	33
SNO+	R	<b>159</b>	217	136	—	—	118
	O	<b>50</b>	69	43	—	—	37
SuperNEMO	R	<b>208</b>	171	139	274	158	174
	O	<b>63</b>	52	42	83	48	53

**Table 4.** Sensitivity of the different proposals at 90% CL in the R and O case, taking into account also the projected isotope mass that can be collected by the different proposals and for a 10 years data-taking period. All sensitivity numbers refer to  $m_{\beta\beta}$  values in meV units. The PMR column relies on our estimate for NME most probable values, while subsequent columns use published NME results (see Sec. 2 for details).

large masses, and reaching the collaboration’s target mass of 100 kg appears as a formidable challenge.

The scalability of SNO+ depends on the feasibility of enriching neodymium, most likely a difficult enterprise. In case of success, they could multiply their mass by a factor 10.

All the xenon-based experiments can, in principle, use one ton of enriched xenon or more for the next-to-next generation. Notice that 700 kg of  $^{136}\text{Xe}$  have already been acquired by the 3 collaborations. Clearly the best isotope among the ones considered here, as far as scalability to large  $\beta\beta$  isotope masses is concerned, appears to be  $^{136}\text{Xe}$ .

Table 4 summarizes our  $m_{\beta\beta}$  sensitivity findings if we further assume the reference and optimistic values for the  $\beta\beta$  isotope masses as given in Tab. 2, and for a data-taking period of 10 years, common to all proposals. In the reference scenario, CUORE, EXO, KamLAND-Zen and NEXT cover most of the degenerate spectrum, with CUORE and KamLAND-Zen reaching the best sensitivity ( $m_{\beta\beta} \simeq 62$  meV at 90% CL assuming our PMR estimate for NME most probable values). In the optimistic scenario, these four experiments cover part of the inverted hierarchy, with NEXT reaching the best sensitivity ( $m_{\beta\beta} \simeq 27$  meV at 90% CL). Even taking into account the uncertainties associated to this calculation, it appears that the scalability to large masses and expected low background of the xenon experiments give them an advantage over other experimental approaches, with the possible exception of CUORE.

For completeness, Tab. 4 also shows how  $m_{\beta\beta}$  sensitivity results are affected by different assumptions regarding NME values. When available, ISM NME values [10, 11] give  $m_{\beta\beta}$  sensitivity results that are about 30% worse than what obtained from our estimate for NME

most probable values (PMR column in Tab. 4). By construction, non-ISM NME values [12–18] tend to give  $m_{\beta\beta}$  sensitivity results that are better than the PMR ones by a similar relative amount.

## 7 Conclusions

The answer to the question of the neutrino particle/antiparticle nature has become a critical issue after the observation of neutrino oscillations. Not only it would improve our understanding of this intriguing fermion, but it would also have an enormous impact in many other areas of fundamental physics.

Neutrinoless double beta decay ( $\beta\beta 0\nu$ ) is possible if and only if neutrinos are massive Majorana particles. The observation of such a process is considered the most promising strategy to demonstrate the Majorana nature of the neutrino. Furthermore, the measurement of the lifetime for this process would provide direct information on the absolute scale of neutrino masses.

An obvious goal for the next-generation experiments is to push the current limits down to Majorana masses  $m_{\beta\beta}$  of at least 100 meV to unambiguously confirm or discard the evidence for  $\beta\beta 0\nu$  claimed by part of the Heidelberg-Moscow Collaboration [7], with sufficient statistics. This appears to be within the range of all the proposals. A more ambitious goal is to fully explore the degenerate spectrum, down to  $m_{\beta\beta} \simeq 50$  meV. None of the proposals appears quite capable of this, under the assumptions discussed in this work (the reference scenario, 90% confidence level, data-taking period of 10 years), although CUORE ( $m_{\beta\beta} \simeq 62$  meV), KamLAND-Zen ( $m_{\beta\beta} \simeq 62$  meV), NEXT ( $m_{\beta\beta} \simeq 71$  meV) and EXO ( $m_{\beta\beta} \simeq 79$  meV) end up quite close.

The optimistic scenario described in the text may be implemented by the next-to-next generation of experiments. The goal here would be to at least partially explore the inverse hierarchy. This seems to be within the reach of NEXT ( $m_{\beta\beta} \simeq 27$  meV), CUORE ( $m_{\beta\beta} \simeq 30$  meV), KamLAND-Zen ( $m_{\beta\beta} \simeq 32$  meV) and perhaps EXO ( $m_{\beta\beta} \simeq 41$  meV).

More generally, the suitability of SNO+, GERDA and SuperNEMO for a next-to-next round of experiments is essentially dependent upon their capability to scale the technology to large masses. Scalability appears easier for CUORE, in particular if the techniques being developed in the context of the LUCIFER Collaboration (see for example [2]) permit large background rate reductions. Finally it appears that  $^{136}\text{Xe}$  would be a particularly favorable isotope to use, since it permits target masses of 1 ton or more and low-background experimental techniques such as the ones proposed by NEXT, KamLAND-Zen and EXO.

Last but not least, while we believe that the general trends observed in the sensitivity comparisons make sense, one should take the numbers presented here *cum grano salis*. There are many uncertainties, ranging from the precise values of the nuclear matrix elements to the ultimate background rate and isotope mass that can be achieved by the different experimental techniques. Nevertheless, we suggest that  $\beta\beta 0\nu$  collaborations adopt the simple and unambiguous procedure we have described in this work to derive  $m_{\beta\beta}$  sensitivities, in order to allow for more meaningful physics case comparisons.

## Acknowledgments

The authors acknowledge support by the Spanish Ministry of Science and Innovation (MICINN) under grants CONSOLIDER-INGENIO 2010 CSD2008-0037 (CUP), FPA2009-13697-C04-04

and RYC-2008-03169. The authors thank Bob Cousins, Fred James, Pilar Hernández and Mauro Mezzetto for illuminating discussions. We also appreciate the comments of several of our colleagues in the double beta decay community, including Alessandro Bettini, Peter Grabmayr, Franco Iachello, Stefan Schönert, Steve Elliott, Giorgio Gratta, and Andrea Giuliani.

## A Construction of unified-approach confidence intervals

For completeness and for pedagogical purposes, in this appendix we describe in some detail how to construct a confidence interval for a signal mean to be inferred from observations, using the unified approach by Feldman and Cousins [8] and in the case of a Poisson process with background. In this case, the relevant pdf to observe  $n$  events given a (unknown) signal mean  $\mu$  and a (known) background mean  $b$  is given by the Poisson distribution:

$$\text{Po}(n; \mu + b) = \frac{(\mu + b)^n}{n!} e^{-(\mu+b)}, \quad (\text{A.1})$$

Just as can be done for classical confidence intervals (that is, the one-sided upper confidence limit and the two-sided central confidence interval discussed in Secs. 4.1 and 4.2), the unified approach uses the Neyman construction of confidence belts [35]. In this construction, given the known mean background expectation  $b$  and for each value of the unknown signal mean  $\mu$ , one selects an interval in  $n$  such that:

$$\sum_{n=n_1}^{n_2} \text{Po}(n; \mu + b) \geq \text{CL} \quad (\text{A.2})$$

where CL is the desired confidence level, for example  $\text{CL} = 0.90$ . Once this exercise is repeated for all possible values of  $\mu$  and once the experiment's outcome  $n_{\text{obs}}$  is known, the confidence interval  $[\mu_{\text{lo}}, \mu_{\text{up}}]$  for the signal mean  $\mu$  can be extracted.

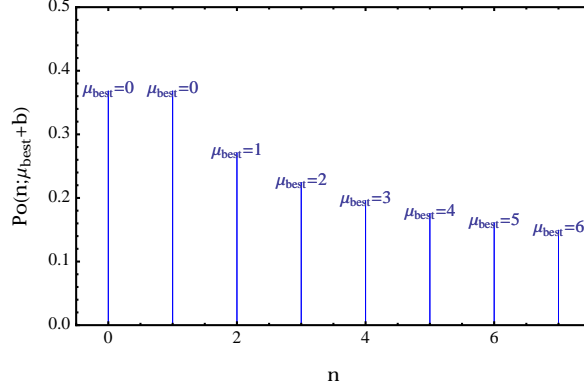
However, while the acceptance interval  $n \in [n_1, n_2]$  is constrained by Eq. A.2, it is not fully specified by it. The freedom on how to specify such intervals marks the difference between the classical and the unified approach confidence belt construction. In the unified approach method, such acceptance intervals at fixed  $\mu$  are determined based on a ordering principle for the  $n$  values based on likelihood ratios, as will be shown below. The entire procedure can be summarized as follows:

1. Compute the mean background expectation,  $b$ . Suppose that, in our case,  $b = 1.0$ . Note that  $b$  does not need to be an integer value, in general.
2. For the given mean background expectation  $b$  and for each possible measurement outcome  $n$ , compute the best estimator  $\mu_{\text{best}}$  for the true value of the mean signal yield  $\mu$ . If  $\mu$  were unconstrained, the best estimator  $\mu_{\text{best}}$  would simply be found by maximizing the Poisson probability for any given  $b$  and  $n$ :

$$\left. \frac{d\text{Po}(n; \mu + b)}{d\mu} \right|_{n,b} = 0 \Rightarrow \mu_{\text{best}} = n - b. \quad (\text{A.3})$$

However, we know that only non-negative values for  $\mu$  are physically allowed, so that our best estimator is given by:

$$\mu_{\text{best}} = \max(0, n - b) \quad (\text{A.4})$$



**Figure 9.** For a given mean background expectation  $b = 1.0$ ,  $\text{Po}(n; \mu_{\text{best}} + b)$  as a function of all possible measurement outcomes  $n$  is shown. For each  $n$ , the value  $\mu_{\text{best}}$  is also shown in the figure,

In the equations above, the measurement outcome  $n$  can take any non-negative integer value, while  $\mu_{\text{best}}$  can take any non-negative value. The values of  $\mu_{\text{best}}$  for  $b = 1.0$  and for each possible measurement outcome  $n$  are reported in Tab. 5.

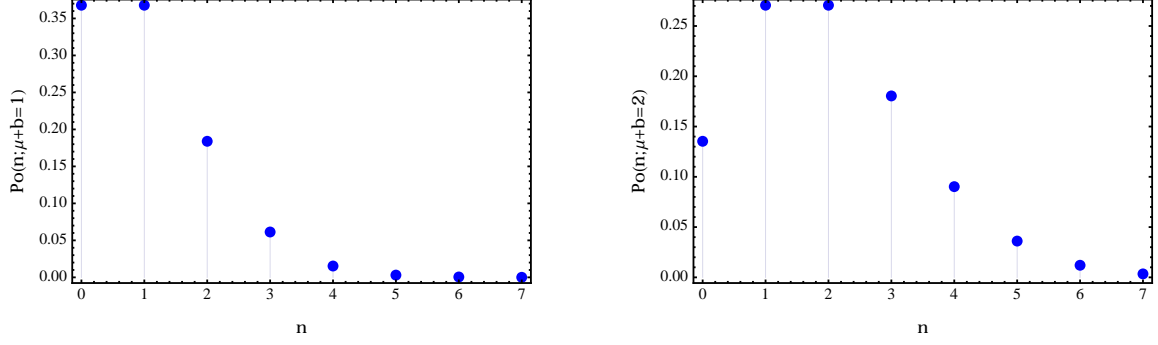
$n$	$\mu_{\text{best}}$	$\text{Po}(n; \mu_{\text{best}} + b)$
0	0.0	0.368
1	0.0	0.368
2	1.0	0.271
3	2.0	0.224
4	3.0	0.195
5	4.0	0.175

**Table 5.** For a given mean background expectation  $b = 1.0$ , values for the signal best estimator  $\mu_{\text{best}}$  and for  $\text{Po}(n; \mu_{\text{best}} + b)$  as a function of all possible measurement outcomes  $n$ .

3. For the given mean background expectation  $b$  and for each possible measurement outcome  $n$ , compute the likelihood  $\text{Po}(n; \mu_{\text{best}} + b)$  of obtaining  $n$  given the best-fit physically allowed signal mean  $\mu_{\text{best}}$ . This is reported in tabular form in Tab. 5, and in graphical form in Fig. 9. Explicitly, the function  $\text{Po}(n; \mu_{\text{best}} + b)$  is equal to:

$$\text{Po}(n; \mu_{\text{best}} + b) = \begin{cases} \frac{b^n}{n!} e^{-b}, & \text{if } n \leq b (\Rightarrow \mu_{\text{best}} = 0) \\ \frac{n^n}{n!} e^{-n}, & \text{if } n > b (\Rightarrow \mu_{\text{best}} > 0) \end{cases} \quad (\text{A.5})$$

4. We now consider a possible value for the unknown true signal mean  $\mu$ . For simplicity, we will show the computational details only for  $\mu = 0$  and for  $\mu = 1.0$ . Note that  $\mu$  does not need to be an integer value, in general.
5. For the given mean background expectation  $b$  (fixed in step 1), and for the given true signal mean  $\mu$  (fixed in step 4), order the possible measurement outcomes  $n$  from most to least likely. This ordering is done according to the values of the likelihood ratio (also



**Figure 10.** The function  $\text{Po}(n; \mu + b)$  for a mean background expectation  $b = 1.0$ , and a signal mean  $\mu = 0$  (left) or  $\mu = 1.0$  (right).

called *profile likelihood*) [71]:

$$\mathcal{L}_R \equiv \frac{\text{Po}(n; \mu + b)}{\text{Po}(n; \mu_{\text{best}} + b)}, \quad (\text{A.6})$$

where the likelihood  $\text{Po}(n; \mu_{\text{best}} + b)$  was computed in step 3, while  $\text{Po}(n; \mu + b)$  is the likelihood of obtaining  $n$  for the given signal mean  $\mu$ . Explicitly, the function  $\mathcal{L}_R$  is equal to:

$$\mathcal{L}_R = \begin{cases} \left(\frac{\mu+b}{b}\right)^n e^{-\mu}, & \text{if } n \leq b \ (\Rightarrow \mu_{\text{best}} = 0) \\ \left(\frac{\mu+b}{n}\right)^n e^{-(\mu+b-n)}, & \text{if } n > b \ (\Rightarrow \mu_{\text{best}} > 0) \end{cases} \quad (\text{A.7})$$

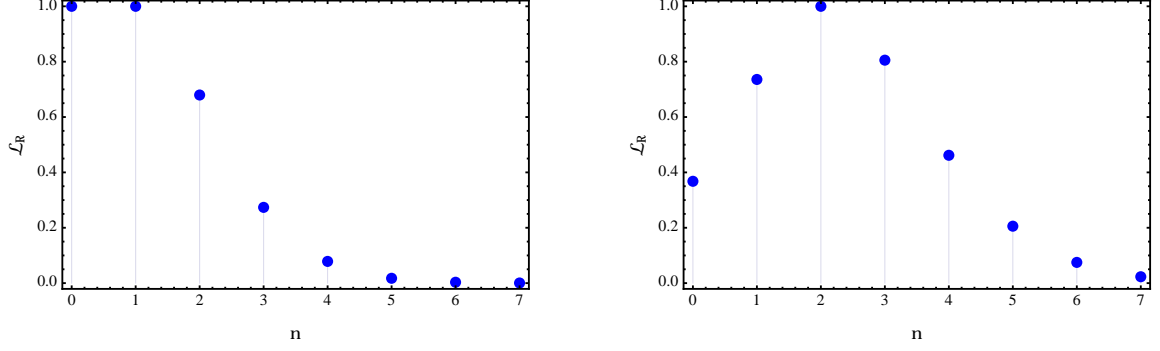
The highest rank ( $\mathcal{R} = 1$ ) is assigned to the  $n$  value having the highest value of  $\mathcal{L}_R$ .

As examples,  $\text{Po}(n; \mu + b)$  values for each  $n$  are given in the left panels of Tab. 6 and of Fig. 10 for  $\mu = 0$ , and in the right panels of Tab. 6 and of Fig. 10 for  $\mu = 1.0$ . The corresponding likelihood ratios  $\mathcal{L}_R$  and ranking of  $n$  values from most to least likely are also given in Tab. 6 for  $\mu = 0$  and 1.0. The likelihood ratio distributions are also shown in graphical form in Fig. 11, again for  $\mu = 0$  (left) and  $\mu = 1.0$  (right). Note that  $\mathcal{L}_R$  is monotonically decreasing with  $n$  for  $\mu = 0$ , but this is not the case for  $\mu = 1.0$ , where  $n = 2$  has the maximum rank.

$n$	$\text{Po}(n; \mu + b)$	$\mathcal{L}_R$	$\mathcal{R}$	$\mathcal{P}$
0	0.368	1.000	1	0.368
1	0.368	1.000	2	0.736
2	0.184	0.680	3	0.920
3	0.061	0.274		
4	0.015	0.078		
5	0.003	0.017		

$n$	$\text{Po}(n; \mu + b)$	$\mathcal{L}_R$	$\mathcal{R}$	$\mathcal{P}$
0	0.135	0.368	5	0.947
1	0.271	0.736	3	0.722
2	0.271	1.000	1	0.271
3	0.180	0.805	2	0.451
4	0.090	0.462	4	0.812
5	0.036	0.206		

**Table 6.** For a given mean background expectation  $b = 1.0$  and mean signal  $\mu = 0$  (left table) and  $\mu = 1.0$  (right table), ranking  $\mathcal{R}$  of the  $n$  values according to the likelihood ratio  $\mathcal{L}_R = \text{Po}(n; \mu + b)/\text{Po}(n; \mu_{\text{best}} + b)$ , and confidence interval construction in  $n$  until a cumulative probability  $\mathcal{P} = 0.90$  is reached. See text for details.

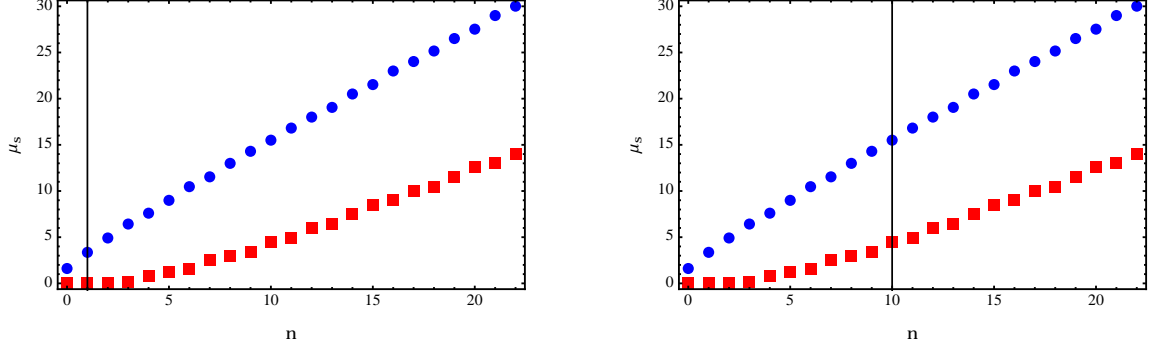


**Figure 11.** The function  $\mathcal{L}_R = \text{Po}(n; \mu + b) / \text{Po}(n; \mu_{\text{best}} + b)$  for  $\mu = 0$  (left) and  $\mu = 1.0$  (right).

6. For a given  $b$  and  $\mu$ , construct a confidence interval in  $n$  by adding the  $\text{Po}(n; \mu + b)$  values until the cumulative probability is at least as large as the desired confidence level. In this confidence interval construction, the  $\text{Po}(n; \mu + b)$  values are added according to their ranking (as defined in step 5), starting from the most likely  $n$  value ( $\mathcal{R} = 1$ ). As examples, Tab. 6 (left) shows that only three  $n$  values are necessary to reach a 90% CL for the  $\mu = 0$  case. The three values are added in this order:  $n = 0, 1, 2$ . The resulting confidence interval is therefore  $n \in [0, 2]$ . Similarly, Tab. 6 (right) shows that five  $n$  values are needed to reach a 90% CL:  $n = 2, 3, 1, 4, 0$ , resulting in the interval  $n \in [0, 4]$ . The larger the desired confidence level is, the larger is the corresponding  $n$  interval.
7. Repeat steps 4-6 for different values of  $\mu$ . In this scan over all possible values of the unknown true signal mean  $\mu$ , one typically starts from  $\mu = 0$ , and gradually increases the  $\mu$  values according to some  $\mu$  step size. Even though we showed the cases  $\mu = 0$  and  $\mu = 1.0$  only, the step size is often chosen to be much smaller than unity when small signals are searched for.

Once the  $n$  confidence intervals have been obtained for each  $\mu$ , the confidence belt construction is complete for the known mean background expectation  $b$  and for the desired confidence level (90% CL, in our case). Note that this confidence belt construction does not require any knowledge of the actual experiment's outcome,  $n_{\text{obs}}$ .

8. Now perform the measurement. Suppose that the result is  $n_{\text{obs}} = 1$ , that is, compatible with the background-only mean expectation  $b = 1.0$ . The left panel in Fig. 12 shows a vertical line corresponding to this value. The intercepts of the confidence belts with the  $n_{\text{obs}} = 1$  vertical line fixes the allowed range in the true signal mean,  $\mu \in [\mu_{\text{lo}}, \mu_{\text{up}}]$ . Given that the observable  $n$  is discrete, one more prescription is needed in this case to fully specify the range  $\mu \in [\mu_{\text{lo}}, \mu_{\text{up}}]$ . When several values of  $\mu$  yield the same acceptance interval  $n \in [n_1, n_2]$  in Eq. A.2, the constant  $n_{\text{obs}}$  vertical line does not intersect the belts in only two points, but rather along two vertical segments. When this occurs, we conservatively take the confidence interval to have  $\mu_{\text{lo}}$  corresponding to the smallest value of  $\mu$  with  $n_2 = n_{\text{obs}}$ , and  $\mu_{\text{up}}$  as the largest value of  $\mu$  with  $n_1 = n_{\text{obs}}$ . As can be read from the left panel in Fig. 12, in this case the range is  $\mu \in [0, 3.4]$ . On the other hand, suppose that the experiment had measured a significant excess above the background-only prediction:  $n_{\text{obs}} = 10$ . The right panel in Fig. 12 shows the  $\mu$  range that would have been obtained in this case:  $\mu \in [4.5, 15.5]$ .



**Figure 12.** Confidence belts constructed for a Poisson process with mean background expectation  $b = 1.0$ , using the unified approach and for a 90% confidence level. The left (right) panel shows how to infer a range in the signal yield  $\mu$  given a measurement outcome of  $n_{\text{obs}} = 1(10)$ .

The unified approach procedure described above solves two problems of classical confidence belts, namely that: (a) a downward fluctuation in background can produce an empty set confidence level (see Sec. 4.2), and (b) using the measured result to decide whether to use a central or upper ordering principle leads to the wrong coverage (*flip-flopping*, see [8]). The unified approach never produces empty confidence intervals, and provides the correct frequentist coverage by avoiding to use the measurement outcome to construct the confidence belts. In this approach, if the lower bound obtained by the above procedure is strictly  $\mu_{\text{lo}} = 0$ , then an upper limit is quoted; otherwise, if  $\mu_{\text{lo}} > 0$ , a central confidence interval is given. With these definitions of upper limit and central interval, and as can be seen in Fig. 12, this method smoothly transitions from an upper limit to a central interval for the signal  $\mu$  as one moves from a null result (compatible with background-only expectation, left panel) to a non-null result (right panel).

## References

- [1] F. T. Avignone III, S. R. Elliott, and J. Engel, *Double Beta Decay, Majorana Neutrinos, and Neutrino Mass*, *Rev. Mod. Phys.* **80** (2008) 481–516, [[arXiv:0708.1033](#)].
- [2] A. Giuliani, *Searches for neutrinoless double beta decay*, *Acta Phys. Polon.* **B41** (2010) 1447–1468.
- [3] J. Schechter and J. W. F. Valle, *Neutrinoless double-beta decay in  $SU(2) \times U(1)$  theories*, *Phys. Rev.* **D25** (1982) 2951.
- [4] M. Hirsch, S. Kovalenko, and I. Schmidt, *Extended Black Box Theorem for Lepton Number and Flavor Violating processes*, *Phys. Lett.* **B642** (2006) 106–110, [[hep-ph/0608207](#)].
- [5] S. R. Elliott and P. Vogel, *Double beta decay*, *Ann. Rev. Nucl. Part. Sci.* **52** (2002) 115–151, [[hep-ph/0202264](#)].
- [6] M. C. González-García and M. Maltoni, *Phenomenology with Massive Neutrinos*, *Phys. Rept.* **460** (2008) 1–129, [[arXiv:0704.1800](#)].
- [7] H. V. Klapdor-Kleingrothaus, A. Dietz, H. L. Harney, and I. V. Krivosheina, *Evidence for Neutrinoless Double Beta Decay*, *Mod. Phys. Lett.* **A16** (2001) 2409–2420, [[hep-ph/0201231](#)].
- [8] G. J. Feldman and R. D. Cousins, *A Unified Approach to the Classical Statistical Analysis of Small Signals*, *Phys. Rev.* **D57** (1998) 3873–3889, [[physics/9711021](#)].

- [9] E. Caurier, J. Menéndez, F. Nowacki, and A. Poves, *The Influence of pairing on the nuclear matrix elements of the neutrinoless beta beta decays*, *Phys. Rev. Lett.* **100** (2008) 052503, [[arXiv:0709.2137](#)].
- [10] J. Menendez, A. Poves, E. Caurier, and F. Nowacki, *Disassembling the Nuclear Matrix Elements of the Neutrinoless beta beta Decay*, *Nucl.Phys.* **A818** (2009) 139–151, [[arXiv:0801.3760](#)].
- [11] J. Menéndez, A. Poves, E. Caurier, and F. Nowacki, *The Occupancies of individual orbits and the nuclear matrix element of the Ge-76 neutrinoless beta beta decay*, *Phys.Rev.* **C80** (2009) 048501, [[arXiv:0906.0179](#)].
- [12] J. Suhonen and O. Civitarese, *Effects of orbital occupancies and spin-orbit partners on  $0\nu\beta\beta$  decay rates*, *Nucl.Phys.* **A847** (2010) 207–232.
- [13] F. Simkovic, A. Faessler, V. Rodin, P. Vogel, and J. Engel, *Anatomy of nuclear matrix elements for neutrinoless double-beta decay*, *Phys.Rev.* **C77** (2008) 045503, [[arXiv:0710.2055](#)].
- [14] F. Simkovic, A. Faessler, H. Muther, V. Rodin, and M. Stauf, *The  $0\nu\beta\beta$ -decay nuclear matrix elements with self-consistent short-range correlations*, *Phys. Rev.* **C79** (2009) 055501, [[arXiv:0902.0331](#)].
- [15] J. Barea and F. Iachello, *Neutrinoless double-beta decay in the microscopic interacting boson model*, *Phys. Rev.* **C79** (2009) 044301.
- [16] T. R. Rodriguez and G. Martinez-Pinedo, *Energy density functional study of nuclear matrix elements for neutrinoless  $\beta\beta$  decay*, *Phys.Rev.Lett.* **105** (2010) 252503, [[arXiv:1008.5260](#)].
- [17] F. Simkovic, A. Faessler, and P. Vogel,  *$0\nu\beta\beta$  nuclear matrix elements and the occupancy of individual orbits*, *Phys.Rev.* **C79** (2009) 015502, [[arXiv:0812.0348](#)].
- [18] D.-L. Fang, A. Faessler, V. Rodin, and F. Simkovic, *Neutrinoless double-beta decay of  $^{150}\text{Nd}$  accounting for deformation*, *Phys.Rev.* **C82** (2010) 051301, [[arXiv:1009.5579](#)].
- [19] M. Blennow, E. Fernández-Martínez, J. López-Pavón, and J. Menéndez, *Neutrinoless double beta decay in seesaw models*, *JHEP* **1007** (2010) 096, [[arXiv:1005.3240](#)].
- [20] J. Menendez, A. Poves, E. Caurier, and F. Nowacki, *Novel Nuclear Structure Aspects of the  $0\nu\beta\beta$  Decay*, *J.Phys.Conf.Ser.* **267** (2011) 012058, [[arXiv:1006.5631](#)].
- [21] H. Feldmeier, T. Neff, R. Roth, and J. Schnack, *A Unitary correlation operator method*, *Nucl.Phys.* **A632** (1998) 61–95, [[nucl-th/9709038](#)].
- [22] F. Simkovic, G. Pantis, J. Vergados, and A. Faessler, *Additional nucleon current contributions to neutrinoless double beta decay*, *Phys.Rev.* **C60** (1999) 055502, [[hep-ph/9905509](#)].
- [23] A. Smolnikov and P. Grabmayr, *Conversion of experimental half-life to effective electron neutrino mass in  $0\nu\beta\beta$  decay*, *Phys.Rev.* **C81** (2010) 028502.
- [24] J. Schiffer, S. Freeman, J. Clark, C. Deibel, C. Fitzpatrick, *et. al.*, *Nuclear Structure Relevant to Neutrinoless Double Beta Decay: Ge-76 and Se-76*, *Phys.Rev.Lett.* **100** (2008) 112501, [[arXiv:0710.0719](#)].
- [25] B. Kay, J. Schiffer, S. Freeman, T. Adachi, J. Clark, *et. al.*, *Nuclear structure relevant to neutrinoless double beta decay: The valence protons in Ge-76 and Se-76*, *Phys.Rev.* **C79** (2009) 021301, [[arXiv:0810.4108](#)].
- [26] G. Audi, A. H. Wapstra, and C. Thibault, *The 2003 atomic mass evaluation: (ii). tables, graphs and references*, *Nuclear Physics A* **729** (2003), no. 1 337.
- [27] B. J. Mount, M. Redshaw, and E. G. Myers, *Double-beta-decay  $Q$  values of Se-74 and Ge-76*, *Phys. Rev.* **C81** (2010) 032501.
- [28] J. N. Nxumalo, J. G. Hykawy, P. P. Unger, C. A. Lander, R. C. Barber, K. S. Sharma, and

- H. E. Duckworth, *Energy available for the double beta decay of  $^{82}\text{Se}$* , *Physics Letters B* **302** (1993), no. 1 13 – 14.
- [29] M. Redshaw, B. J. Mount, E. G. Myers, and F. T. Avignone, *Masses of Te-130 and Xe-130 and Double-beta-Decay Q Value of Te-130*, *Phys. Rev. Lett.* **102** (2009) 212502.
- [30] M. Redshaw, E. Wingfield, J. McDaniel, and E. G. Myers, *Mass and double-beta-decay Q value of Xe-136*, *Phys. Rev. Lett.* **98** (2007) 053003.
- [31] V. S. Kolhinen *et. al.*, *Double-beta decay Q value of Nd-150*, *Phys. Rev.* **C82** (2010) 022501.
- [32] K. Zuber, *Summary of the workshop on: 'Nuclear matrix elements for neutrinoless double beta decay'*, [nucl-ex/0511009](#).
- [33] H. Klapdor-Kleingrothaus *et. al.*, *Latest results from the Heidelberg-Moscow double beta decay experiment*, *Eur. Phys. J.* **A12** (2001) 147–154, [[hep-ph/0103062](#)].
- [34] C. E. Aalseth *et. al.*, *Comment on 'Evidence for Neutrinoless Double Beta Decay'*, *Mod. Phys. Lett.* **A17** (2002) 1475–1478, [[hep-ex/0202018](#)].
- [35] J. Neyman, *Outline of a Theory of Statistical Estimation Based on the Classical Theory of Probability*, *Phil. Trans. R. Soc. A* **236** (1937), no. 767 333–380.
- [36] F. James, (Ed. ), Y. Perrin, (Ed. ), and L. Lyons, (Ed. ), *Workshop on confidence limits, CERN, Geneva, Switzerland, 17-18 Jan 2000: Proceedings*, . CERN-2000-005.
- [37] **GERDA** Collaboration, I. Abt *et. al.*, *A new Ge-76 double beta decay experiment at LNGS*, [hep-ex/0404039](#).
- [38] M. Barnabé Heider *et. al.*, *Operation of a GERDA Phase I prototype detector in liquid argon and nitrogen*, [arXiv:0812.3976](#).
- [39] I. Barabanov *et. al.*, *Shielding of the GERDA experiment against external gamma background*, *Nucl. Instrum. Meth.* **A606** (2009) 790–794.
- [40] **GERDA** Collaboration, G. Zuzel, *The GERDA experiment at Gran Sasso*, *Acta Phys. Polon.* **B41** (2010) 1469.
- [41] **MAJORANA** Collaboration, S. Elliott *et. al.*, *The Majorana Project*, *J.Phys.Conf.Ser.* **173** (2009) 012007, [[arXiv:0807.1741](#)].
- [42] **MAJORANA** Collaboration, V. Guiseppe *et. al.*, *The Majorana Neutrinoless Double-Beta Decay Experiment*, [arXiv:0811.2446](#).
- [43] **MAJORANA** Collaboration, R. Henning *et. al.*, *The MAJORANA DEMONSTRATOR: An R&D project towards a tonne-scale germanium neutrinoless double-beta decay search*, *AIP Conf.Proc.* **1182** (2009) 88–91, [[arXiv:0907.1581](#)].
- [44] **CUORE** Collaboration, C. Arnaboldi *et. al.*, *CUORE: A cryogenic underground observatory for rare events*, *Nucl. Instrum. Meth.* **A518** (2004) 775–798, [[hep-ex/0212053](#)].
- [45] **CUORE** Collaboration, M. Sisti, *From Cuoricino to CUORE: Investigating neutrino properties with double beta decay*, *J. Phys. Conf. Ser.* **203** (2010) 012069.
- [46] F. Bellini *et. al.*, *Monte Carlo evaluation of the external gamma, neutron and muon induced background sources in the CUORE experiment*, *Astropart. Phys.* **33** (2010) 169–174, [[arXiv:0912.0452](#)].
- [47] **EXO** Collaboration, N. Ackerman, *Status of EXO-200*, [arXiv:0909.1826](#).
- [48] **EXO** Collaboration, R. Gornea, *Double beta decay in liquid xenon*, *J. Phys. Conf. Ser.* **179** (2009) 012004.
- [49] **EXO** Collaboration, E. Conti *et. al.*, *Correlated fluctuations between luminescence and ionization in liquid xenon*, *Phys. Rev.* **B68** (2003) 054201, [[hep-ex/0303008](#)].

- [50] M. Danilov *et. al.*, *Detection of very small neutrino masses in double-beta decay using laser tagging*, *Phys. Lett.* **B480** (2000) 12–18, [[hep-ex/0002003](#)].
- [51] **NEXT** Collaboration, F. Grañena *et. al.*, *NEXT, a HPGXe TPC for neutrinoless double beta decay searches*, [arXiv:0907.4054](#).
- [52] D. Nygren, *High-pressure xenon gas electroluminescent TPC for Onu beta beta-decay search*, *Nucl. Instrum. Meth.* **A603** (2009) 337–348.
- [53] **SNO+** Collaboration, M. C. Chen, *The SNO+ Experiment*, [arXiv:0810.3694](#).
- [54] A. Wright, *Robust Signal Extraction Methods and Monte Carlo Sensitivity Studies for the Sudbury Neutrino Observatory and SNO+ Experiments*. PhD thesis, Queen’s University, 2009.
- [55] **SNO+** Collaboration, C. Kraus and S. J. M. Peeters, *The rich neutrino programme of the SNO+ experiment*, *Prog. Part. Nucl. Phys.* **64** (2010) 273–277.
- [56] **KamLAND** Collaboration, A. Terashima *et. al.*, *R&D for possible future improvements of KamLAND*, *J. Phys. Conf. Ser.* **120** (2008) 052029.
- [57] M. Koga, *KamLAND double beta decay experiment using  $^{136}\text{Xe}$* , in *International Conference on High Energy Physics (ICHEP)*, (Paris), July, 2010.
- [58] **NEMO/SuperNEMO** Collaboration, R. Saakyan, *Topological detection of double beta decay with NEMO3 and SuperNEMO*, *J. Phys. Conf. Ser.* **179** (2009) 012006.
- [59] **NEMO/SuperNEMO** Collaboration, Y. A. Shitov, *A Search for Neutrinoless Double Beta Decay: from NEMO-3 to SuperNEMO*, [arXiv:1006.4775](#).
- [60] P. Novella, *Experimental studies of neutrino nature*. PhD thesis, Universidad de Valencia, 2008.
- [61] **NEMO** Collaboration, R. Arnold *et. al.*, *First results of the search of neutrinoless double beta decay with the NEMO 3 detector*, *Phys. Rev. Lett.* **95** (2005) 182302, [[hep-ex/0507083](#)].
- [62] **NEMO** Collaboration, J. Argyriades *et. al.*, *Measurement of the Double Beta Decay Half-life of 150-Nd and Search for Neutrinoless Decay Modes with the NEMO-3 Detector*, *Phys. Rev.* **C80** (2009) 032501, [[arXiv:0810.0248](#)].
- [63] **NEMO/SuperNEMO** Collaboration, J. Argyriades *et. al.*, *Measurement of the background in the NEMO 3 double beta decay experiment*, *Nucl. Instrum. Meth.* **A606** (2009) 449–465, [[arXiv:0903.2277](#)].
- [64] **SuperNEMO** Collaboration, R. Arnold *et. al.*, *Probing New Physics Models of Neutrinoless Double Beta Decay with SuperNEMO*, [arXiv:1005.1241](#).
- [65] A. Giuliani, *Highlights in double beta decay*, in *International Student Workshop on Neutrinoless Double Beta Decay*, (Laboratori Nazionali del Gran Sasso, Italy), November, 2010.
- [66] P. Barbeau, J. Collar, and O. Tench, *Large-Mass Ultra-Low Noise Germanium Detectors: Performance and Applications in Neutrino and Astroparticle Physics*, *JCAP* **0709** (2007) 009, [[nucl-ex/0701012](#)].
- [67] D. Budjas, M. Barnabe Heider, O. Chkvorets, N. Khanbekov, and S. Schonert, *Pulse shape discrimination studies with a Broad-Energy Germanium detector for signal identification and background suppression in the GERDA double beta decay experiment*, *JINST* **4** (2009) P10007, [[arXiv:0909.4044](#)].
- [68] M. Agostini, C. Ur, D. Budjas, E. Bellotti, R. Brugnera, *et. al.*, *Signal modeling of high-purity Ge detectors with a small read-out electrode and application to neutrinoless double beta decay search in Ge-76*, *JINST* **6** (2011) P03005, [[arXiv:1012.4300](#)].
- [69] **CUORICINO** Collaboration, C. Arnaboldi *et. al.*, *Results from a search for the  $0\nu\beta\beta$ -decay of  $^{130}\text{Te}$* , *Phys. Rev.* **C78** (2008) 035502, [[arXiv:0802.3439](#)].
- [70] R. Luscher *et. al.*, *Search for beta beta decay in Xe-136: New results from the Gotthard*

*experiment*, *Phys. Lett.* **B434** (1998) 407–414.

- [71] W. A. Rolke, A. M. Lopez, and J. Conrad, *Confidence Intervals with Frequentist Treatment of Statistical and Systematic Uncertainties*, *Nucl. Instrum. Meth.* **A551** (2005) 493–503, [[physics/0403059](#)].